
JMIR Medical Informatics

Impact Factor (2023): 3.1

Volume 7 (2019), Issue 1 ISSN 2291-9694 Editor in Chief: Christian Lovis, MD, MPH, FACMI

Contents

Review

- Telemonitoring to Manage Chronic Obstructive Pulmonary Disease: Systematic Literature Review ([e11496](#))
Clemens Kruse, Brandon Pesek, Megan Anderson, Kacey Brennan, Hilary Comfort. 4

Original Papers

- Improving Digital Hospital Transformation: Development of an Outcomes-Based Infrastructure Maturity Assessment Framework ([e12465](#))
Patricia Williams, Brendan Lovelock, Tony Cabarrus, Marlon Harvey. 13
- The Impact of an Electronic Patient Bedside Observation and Handover System on Clinical Practice: Mixed-Methods Evaluation ([e11678](#))
Alexandra Lang, Mark Simmonds, James Pinchin, Sarah Sharples, Lorraine Dunn, Susan Clarke, Owen Bennett, Sally Wood, Caron Swinscoe.
2 8
- Understanding Determinants of Health Care Professionals' Perspectives on Mobile Health Continuance and Performance ([e12350](#))
Ju-Ling Hsiao, Rai-Fu Chen. 41
- European Hospitals' Transition Toward Fully Electronic-Based Systems: Do Information Technology Security and Privacy Practices Follow? ([e11211](#))
Sylvestre Uwizeyemungu, Placide Poba-Nzaou, Michael Cantinotti. 57
- Data Analysis and Visualization of Newspaper Articles on Thirdhand Smoke: A Topic Modeling Approach ([e12414](#))
Qian Liu, Qiuyi Chen, Jiayi Shen, Huailiang Wu, Yimeng Sun, Wai-Kit Ming. 73
- Detection of Bleeding Events in Electronic Health Record Notes Using Convolutional Neural Network Models Enhanced With Recurrent Neural Network Autoencoders: Deep Learning Approach ([e10788](#))
Rumeng Li, Baotian Hu, Feifan Liu, Weisong Liu, Francesca Cunningham, David McManus, Hong Yu. 82
- Use of Electronic Health Record Access and Audit Logs to Identify Physician Actions Following Noninterruptive Alert Opening: Descriptive Study ([e12650](#))
Azraa Amroze, Terry Field, Hassan Fouayzi, Devi Sundaresan, Laura Burns, Lawrence Garber, Rajani Sadasivam, Kathleen Mazor, Jerry Gurwitz, Sarah Cutrona. 92
- Completeness of Electronic Dental Records in a Student Clinic: Retrospective Analysis ([e13008](#))
Seth Levitin, John Grbic, Joseph Finkelstein. 106

OpenEHR and General Data Protection Regulation: Evaluation of Principles and Requirements (e9845)		
Duarte Gonçalves-Ferreira, Mariana Sousa, Gustavo Bacelar-Silva, Samuel Frade, Luís Antunes, Thomas Beale, Ricardo Cruz-Correia.		
1	7	
Transcription Errors of Blood Glucose Values and Insulin Errors in an Intensive Care Unit: Secondary Data Analysis Toward Electronic Medical Record-Glucometer Interoperability (e11873)		
Azizeh Sowan, Ana Vera, Ashwin Malshe, Charles Reed.	129	
Extraction of Geriatric Syndromes From Electronic Health Record Clinical Notes: Assessment of Statistical Natural Language Processing Methods (e13039)		
Tao Chen, Mark Dredze, Jonathan Weiner, Leilani Hernandez, Joe Kimura, Hadi Kharrazi.	138	
Implementation of National Health Informatization in China: Survey About the Status Quo (e12238)		
Chen Li, Xiangdong Xu, Guanghua Zhou, Kai He, Tianliang Qi, Wei Zhang, Feng Tian, Qinghua Zheng, Jianping Hu.	150	
An Assessment of the Interoperability of Electronic Health Record Exchanges Among Hospitals and Clinics in Taiwan (e12630)		
Hsyien-Chia Wen, Wei-Pin Chang, Min-Huei Hsu, Cheng-Hsun Ho, Chi-Ming Chu.	159	
SNOMED CT Concept Hierarchies for Computable Clinical Phenotypes From Electronic Health Record Data: Comparison of Intensional Versus Extensional Value Sets (e11487)		
Ling Chu, Vaishnavi Kannan, Mujeeb Basit, Diane Schaefflein, Adolfo Ortuzar, Jimmie Glorioso, Joel Buchanan, Duwayne Willett.	168	
Utilization of the Veterans Affairs' Transgender E-consultation Program by Health Care Providers: Mixed-Methods Study (e11695)		
John Blosnich, Keri Rodriguez, Kristina Hruska, Dio Kavalieratos, Adam Gordon, Alexis Matza, Susan Mejia, Jillian Shipherd, Michael Kauth.		
1	8	3
Information Technology–Assisted Treatment Planning and Performance Assessment for Severe Thalassemia Care in Low- and Middle-Income Countries: Observational Study (e9291)		
Rajat Agarwal, Amit Sedai, Kumari Ankita, Lalith Parmar, Rakesh Dhanya, Sunil Dhimal, Reshma Srinivas, Ashwini Gowda, Pooja Gujjal, Pushpa H, Suman Jain, J Ramaiah, Sujata Jali, Neelavva Tallur, Stalin Ramprakash, Lawrence Faulkner.	197	
A New Insight Into Missing Data in Intensive Care Unit Patient Profiles: Observational Study (e11605)		
Anis Sharafoddini, Joel Dubin, David Maslove, Joon Lee.	208	
The Connected Intensive Care Unit Patient: Exploratory Analyses and Cohort Discovery From a Critical Care Telemedicine Database (e13006)		
Patrick Essay, Tala Shahin, Baran Balkan, Jarrod Mosier, Vignesh Subbian.	227	
Predicting Appropriate Hospital Admission of Emergency Department Patients with Bronchiolitis: Secondary Analysis (e12591)		
Gang Luo, Bryan Stone, Flory Nkoy, Shan He, Michael Johnson.	240	
Development of Prediction Models Using Machine Learning Algorithms for Girls with Suspected Central Precocious Puberty: Retrospective Study (e11728)		
Liyan Pan, Guangjian Liu, Xiaojian Mao, Huixian Li, Jiexin Zhang, Huiying Liang, Xiuzhen Li.	255	
Medication Use for Childhood Pneumonia at a Children's Hospital in Shanghai, China: Analysis of Pattern Mining Algorithms (e12577)		
Chunlei Tang, Huajun Sun, Yun Xiong, Jiahong Yang, Christopher Vitale, Lu Ruan, Angela Ai, Guangjun Yu, Jing Ma, David Bates.	268	



Corrigenda and Addenda

Figure Correction: Health Information Technology in Healthcare Quality and Patient Safety: Literature Review ([e11320](#))

Sue Feldman, Scott Buchalter, Leslie Hayes. 279

Review

Telemonitoring to Manage Chronic Obstructive Pulmonary Disease: Systematic Literature Review

Clemens Kruse¹, MBA, MSIT, MHA, PhD; Brandon Pesek^{1*}, BA, MHA; Megan Anderson^{1*}, BA, MHA; Kacey Brennan^{1*}, BA, MHA; Hilary Comfort^{1*}, BA, MHA

School of Health Administration, Texas State University, San Marcos, TX, United States

* these authors contributed equally

Corresponding Author:

Clemens Kruse, MBA, MSIT, MHA, PhD

School of Health Administration

Texas State University

601 University Drive

Encino Hall, Room 250

San Marcos, TX, 78666

United States

Phone: 1 512 245 4462

Fax: 1 512 245 8712

Email: scottkruse@txstate.edu

Abstract

Background: Chronic obstructive pulmonary disease (COPD) is a leading cause of death throughout the world. Telemedicine has been utilized for many diseases and its prevalence is increasing in the United States. Telemonitoring of patients with COPD has the potential to help patients manage disease and predict exacerbations.

Objective: The objective of this review is to evaluate the effectiveness of telemonitoring to manage COPD. Researchers want to determine how telemonitoring has been used to observe COPD and we are hoping this will lead to more research in telemonitoring of this disease.

Methods: This review was conducted in accordance with the Assessment for Multiple Systematic Reviews (AMSTAR) and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Authors performed a systematic review of the PubMed and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases to obtain relevant articles. Articles were then accepted or rejected by group consensus. Each article was read and authors identified barriers and facilitators to effectiveness of telemonitoring of COPD.

Results: Results indicate that conflicting information exists for the effectiveness of telemonitoring of patients with COPD. Primarily, 13 out of 29 (45%) articles stated that patient outcomes were improved overall with telemonitoring, while 11 of 29 (38%) indicated no improvement. Authors identified the following facilitators: reduced need for in-person visits, better disease management, and bolstered patient-provider relationship. Important barriers included low-quality data, increased workload for providers, and cost.

Conclusions: The high variability between the articles and the ways they provided telemonitoring services created conflicting results from the literature review. Future research should emphasize standardization of telemonitoring services and predictability of exacerbations.

(*JMIR Med Inform* 2019;7(1):e11496) doi:[10.2196/11496](https://doi.org/10.2196/11496)

KEYWORDS

telemedicine; COPD; chronic disease

Introduction

Rationale

The most recent estimate of the world prevalence for chronic obstructive pulmonary disease (COPD) is 64 million, with 3 million deaths from the disease in 2015 alone [1]. The World Health Organization (WHO) estimates that COPD will be the third-leading cause of death in the world by 2030 and that 90% of its victims live in middle-to-low-income countries [1]. It is primarily caused by cigarette smoke—primary or secondary—and exacerbated by long-term asthma [1]. The United States addressed the increase in prevalence by penalizing reimbursement for public health beneficiaries if a hospital readmitted the patient for the condition within 30 days [2].

The United States also passed the Health Information Technology for Economic and Clinical Health (HITECH) Act in 2009, which incentivized the adoption of health information technology up until 2015 and penalized the lack of adoption thereafter [3]. The HITECH Act served as a catalyst for the diffusion of telemedicine in the United States, which is important because the United States lagged behind other western nations in the use and acceptance of telemedicine. There are many facets to telemedicine, but we will start with a general definition.

The WHO defines telemedicine as follows [4]:

The delivery of health care services, where distance is a critical factor, by all health care professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of health care providers, all in the interests of advancing the health of individuals and their communities.

We choose this definition for our review and, also following the WHO's example, we do not distinguish between telemedicine and telehealth.

US national attention on telemedicine services for rural and other low-access populations has steadily increased in the past decade. As the United States continues to determine how best to fund these services and how to legislate accreditation across state borders and specialties, much research is being conducted on the efficacy of various telemedicine services [5]. While the bulk of US research attends to clinical interventions provided for mental health and chronic diseases, chronic diseases also require regular monitoring of health parameters.

Telemedicine, in its modern form, developed through rapid advancement in communication technology and innovation on the part of health care professionals [6]. Naturally, physicians treating chronic diseases, such as COPD, required methods to track patient health factors and telemonitoring was the solution. Telemonitoring is defined as the distance monitoring of components of a patient's health as part of a larger chronic care model [7]. These methods, when applied to patients with COPD, can utilize caregiver review of data to assess disease state and

health status [8]. Telemonitoring of COPD even has the potential to predict exacerbations before onset [9].

Objective

The objective of this review is to evaluate the effectiveness of telemonitoring to manage the chronic disease of COPD. We want to look at how telemonitoring has been used to observe COPD and we are hoping this will lead to more research in telemonitoring of this disease. We used techniques from the Assessment for Multiple Systematic Reviews (AMSTAR) and reported our findings in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [10,11].

Methods

Stakeholder Involvement in the Review

Neither patients, service users, caregivers, nor lay people were used in the design or execution of this review. The development of outcome measures was not informed by patients' priorities, experience, or preferences. Neither patients, caregivers, nor lay people were involved in the recruitment to and conduct of this review. Since there were no study participants as part of this review, dissemination of results to participants was unnecessary. The development of the research question and outcome measure was not informed by patients' priorities, experience, or preferences. Because there were no study participants, ethics approval and consent to participate were not necessary.

Protocol, Registration, and Information Sources

This systematic literature review followed standard retrieval methods to obtain peer-reviewed articles: multiple-reviewer technique set by the AMSTAR standard to evaluate them and the PRISMA standard to report the analysis conducted on the articles in the review [12,13]. Authors queried the following databases: PubMed and the Cumulative Index to Nursing and Allied Health Literature (CINAHL), which is managed by EBSCO. Authors used the search terms *telemonitoring* and *COPD* and all associated Medical Subject Headings (MeSH) terms: *Chronic Obstructive Pulmonary Disease*, *COAD*, *Chronic Obstructive Airway Disease*, *Chronic Obstructive Lung Disease*, *Airflow Obstruction*, *Chronic or Airflow Obstructions*, *Chronic or Chronic Airflow Obstruction*, and *Chronic Airflow Obstruction*. This review was not registered.

Study Selection and Data Collection Process

Databases were searched for articles published during the time frame of February 1, 2011, through February 1, 2017. Originally, we planned to limit our search to a 5-year span for analysis because of the rapid advancement of technology, but this did not yield a suitable number of articles to analyze. As a result, we expanded our search to a 6-year span. Boolean operators were used during searches to obtain the desired search parameters.

The initial search in PubMed, *telemonitoring* AND *COPD*, returned 88 articles. Restricting the articles further by date eliminated 12 articles; limiting by academic journals and English-only articles eliminated 21 more articles. The initial search in CINAHL generated a total of 38 articles. Restricting

the publication date range minimized this number to a total of 35 articles and removing any articles that were not from academic journals and not written in English resulted in 16 remaining articles.

Data Items and Eligibility Criteria

Authors created a literature matrix detailing the title, author, year, journal, and other pertinent information of the 61 articles in preparation for final screening. To eliminate the possibility for bias, two authors read each abstract and came to a consensus regarding whether the article was germane to the topic. Reviewers agreed that for an article to be accepted for analysis, it had to be published in the last 6 years in a peer-reviewed journal and it had to provide substantive data on the use of telemonitoring to manage COPD. Once all abstracts had been screened for suitability, the authors used a consensus meeting to make final determination on whether to eliminate articles that were not germane to the specific topic and remove any duplicate articles. Through this process, the authors noticed a common reference that had not been caught in their query of the databases. The final number of articles gathered from PubMed and CINAHL totaled 21 and 8, respectively, bringing the final combined total to 29 articles for analysis in the literature review. A kappa statistic was calculated to determine interrater reliability (.91), which is indicative of strong agreement [14,15].

Authors read each article closely and made independent notes of common themes related to each review’s objective. These were shared at a second consensus meeting; through a brief discussion of content and findings, detailed notes were highlighted about barriers and facilitators to the adoption of telemonitoring for the management of COPD. Frequency of occurrence of each of the barriers and facilitators were captured in affinity matrices for further analysis. Data and calculations are available upon request.

Results

Study Selection and Study Characteristics

From the original 136 articles resulting from the initial search, 97 were screened out due to date of publication, nature of publication, and whether the topic was germane to our research (ie, possibly indexed improperly). The list of germane studies was narrowed down to 29. The literature search process is listed in Figure 1.

Results of Individual Studies

The results were mixed regarding the efficacy of telemonitoring to reduce complications associated with COPD. Any clear positive relationship with the use of telemonitoring to manage COPD was obscured. A list of all 29 studies [8,9,16-42] and a summary of each topic is listed in Table 1.

Figure 1. Literature search and selection process.

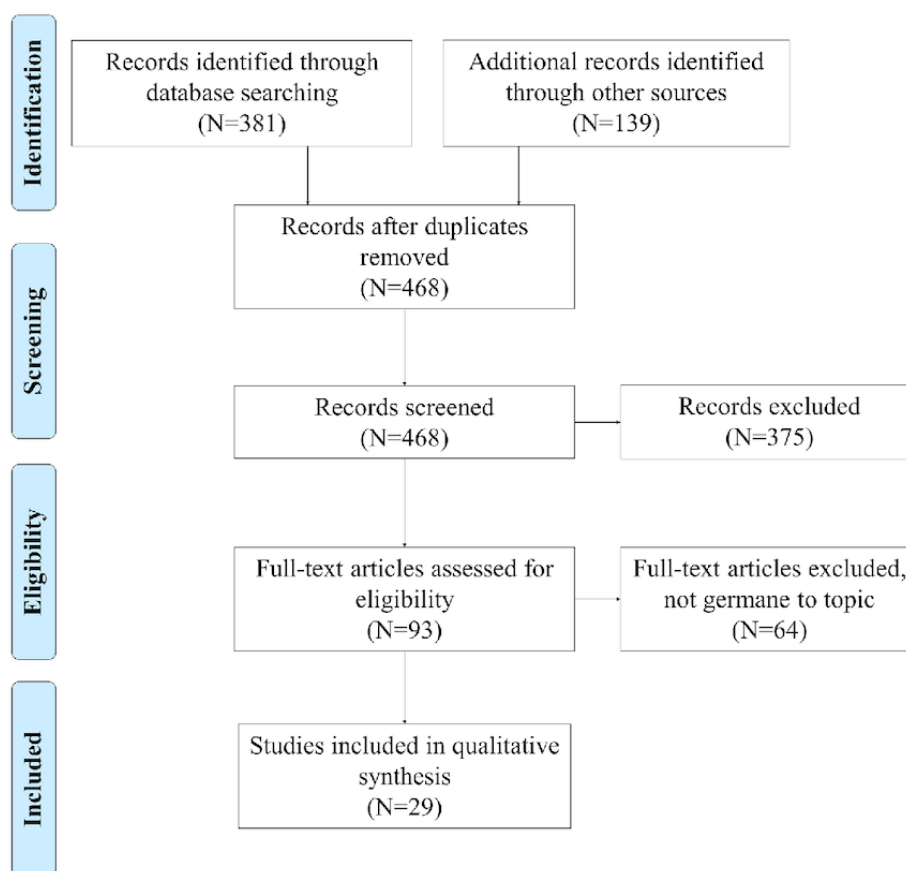


Table 1. Summary of articles analyzed and applicable observations.

Authors and reference	Summary of topic
Jensen MH et al [8]	Clinical impact of home telemonitoring on patients with COPD ^a
Sanchez-Morillo D et al [9]	Pilot study detecting COPD exacerbations early using daily telemonitoring of symptoms and k-means clustering
Alrajab S et al [16]	A home telemonitoring program reduced exacerbation and health care utilization rates in COPD patients with frequent worsening of symptoms
Antoniades NC et al [17]	Pilot study of remote telemonitoring in COPD
Burton C et al [18]	Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease
Celler BG and Sparks RS [19]	Home telemonitoring of vital signs—technical challenges and future directions
Chatwin M et al [20]	Randomized crossover trial of telemonitoring in chronic respiratory patients (TeleCRAFT trial)
Davis C et al [21]	Feasibility and acute care utilization outcomes of a post-acute transitional telemonitoring program for underserved chronic disease patients
Elwyn G et al [22]	Detecting deterioration in patients with chronic disease using telemonitoring: navigating the “trough of disillusionment”
Fairbrother P et al [23]	Exploring telemonitoring and self-management by patients with chronic obstructive pulmonary disease: a qualitative study embedded in a randomized controlled trial
Fairbrother P et al [34]	Continuity, but at what cost? The impact of telemonitoring COPD on continuities of care: a qualitative study
Fernandez-Granero MA et al [25]	Automatic prediction of chronic obstructive pulmonary disease exacerbations through home telemonitoring of symptoms
Fernandez-Granero MA et al [26]	Computerized analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD
Goldstein RS and O'Hoski S [27]	Telemedicine in COPD: time to pause
Ho TW et al [28]	Effectiveness of telemonitoring in patients with chronic obstructive pulmonary disease in Taiwan: a randomized controlled trial
Jordan R et al [29]	Telemonitoring for patients with COPD
Kim J et al [30]	Acceptability of the consumer-centric uHealth services for patients with chronic obstructive pulmonary disease
Kim J et al [31]	Effects of consumer-centered uHealth service for the knowledge, skill, and attitude of the patients with chronic obstructive pulmonary disease
Martin-Lesende I et al [32]	Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on health care resource use (ie, the TELBIL study randomized controlled trial)
McDowell JE et al [33]	A randomized clinical trial of the effectiveness of home-based health care with telemonitoring in patients with COPD
McKinstry B [34]	The use of remote monitoring technologies in managing chronic obstructive pulmonary disease
Pedone C et al [35]	Efficacy of multiparametric telemonitoring on respiratory outcomes in elderly people with COPD: a randomized controlled trial
Pedone C and Lelli D [36]	Systematic review of telemonitoring in COPD: an update
Pinnock H et al [37]	Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: researcher-blind, multicenter, controlled trial
Reddel HK et al [38]	Self-management support and other alternatives to reduce the burden of asthma and chronic obstructive pulmonary disease
Stoddart A et al [39]	Telemonitoring for chronic obstructive pulmonary disease: a cost and cost-utility analysis of a randomized controlled trial
Venter A et al [40]	Results of a telehealth-enabled chronic care management service to support people with long-term conditions at home
Vianello A et al [41]	Home telemonitoring for patients with acute exacerbation of chronic obstructive pulmonary disease: a randomized controlled trial
Zanaboni P et al [42]	Long-term telerehabilitation of COPD patients in their homes: interim results from a pilot study in Northern Norway

^aCOPD: chronic obstructive pulmonary disease.

Table 2. Facilitators to the adoption of telemedicine to manage chronic obstructive pulmonary disease (COPD).

Facilitators	Articles	Occurrence (N=56), n (%)
Improved patient outcomes or satisfaction	[8,9,16,17,24,29-31,36,38,40]	13 (23)
Reduced need for in-person visits	[8,9,16,17,22,30,32,35,36]	9 (16)
Better disease management	[8,17,22,28,30,31,34,40]	8 (14)
Bolstered patient-provider relationship	[20,21,23,24,30]	5 (9)
High-quality data	[16,19,25,28]	4 (7)
Patient empowerment	[8,17,21,23]	4 (7)
Ease of use	[17,25,32]	3 (5)
Predictability of exacerbations	[9,19,25]	3 (5)
Provision of additional services	[30,32,41]	3 (5)
Patient engagement	[16,17]	2 (4)
Access to patient data	[16]	1 (2)
Communication	[22]	1 (2)

Table 3. Barriers to the adoption of telemedicine to manage chronic obstructive pulmonary disease (COPD).

Barriers	Articles	Occurrence (N=57), n (%)
Reduced patient outcomes or no improvement	[16-18,20,28,31,38-42]	11 (19)
Low-quality or limited data	[8,16-18,22-25,28]	9 (16)
Increased workload for providers	[20,19,24,30,32,35,37]	7 (12)
Cost	[29-31,34,39]	5 (9)
Heterogeneity of care	[6,8,21,34,42]	5 (9)
Lack of service standardization	[18,21,22,24,28]	5 (9)
Exacerbations are highly variable	[9,16,18,34]	4 (7)
Uncomfortable with technology	[22,31,33]	3 (5)
Less patient autonomy	[9,20,21]	3 (5)
Time-consuming	[22,24]	2 (4)
Staff shortages or overworked staff	[8]	1 (2)
User perception or perceived lack of usefulness	[23]	1 (2)
User or patient resistance	[24]	1 (2)

Our second consensus meeting helped us identify the 12 facilitators and 13 barriers to the acceptance and feasibility of telemonitoring to manage COPD, which are summarized in [Tables 2](#) and [3](#), respectively.

Facilitators and barriers are sorted by frequency of occurrence. We do not suggest that frequency equates to importance; we highlight only the probability that each theme occurred in the literature.

Synthesis of Results and Risk of Bias Across Studies

Throughout this review, the value of telemonitoring to manage COPD symptoms has been intensively evaluated. Authors identified key facilitators and barriers related to the effectiveness of telemonitoring. The prevalence of factors can be reviewed in [Tables 2](#) and [3](#). Conflicting data were found detailing the efficacy of telemonitoring services for managing COPD. Some articles cited improvements in patient outcomes, satisfaction, anxiety and depression, and hospitalization rates in the facilitator

improved patient outcomes or satisfaction [8,9,16,17,24,29-31,33,34,36,38,40], while others stated that no significant improvement occurred under the barrier *reduced patient outcomes or no improvement* [16-18,20,28,31,38-42].

Articles discuss various causes for improvement of the COPD disease state or perceptions of the disease state. A total of 31% (9/29) of articles stated that telemonitoring reduced the number of in-patient visits required for patients engaged in telemonitoring care, including primary care visits and emergency department visits [8,9,16,17,22,30,32,35,36]. Pinnock [37] and Venter [40] found that enhanced access to care was especially useful in rural areas where access to care may be greatly restricted. As a dominant facilitator for telemonitoring of COPD, the review suggests that telemonitoring interventions have the potential to achieve the main goal of telemedicine services.

A common reason for patient improvements included that providing telemonitoring services to traditional COPD

management added underlying services lines to patient resources [32,34,41]. As more service options were added, including videoconferencing and phone support, articles noted reductions in admissions related to exacerbations. McKinstry [34] found that more successful programs were associated with service lines that were unavailable to regular COPD management programs. Constant access to a respiratory nurse should logically increase patient education and outcomes. Regardless of the number of added service lines provided through telemonitoring of COPD, patients were regularly satisfied with the telemonitoring services provided. Other facilitators, such as higher-quality patient data and ease of use, provide better self-management for patients and more information to caregivers.

Conversely, numerous articles in this review also mentioned the inability for COPD telemonitoring to provide added value for patients [16-18,20,28,31,38-42]. Some of these articles referenced that sample selection did not allow for clear improvements; authors, in some instances, selected patients with excellent self-management practices [36]. Authors also reported that as telemonitoring services expanded, clinician and nurse workloads increased [20,23,24,32,34,37,39]. As staffing is one of the most expensive parts of providing health care, increasing the amount of work required to care for patients can potentially increase costs.

Cost also factored into some of the studies examined [33-35,38,39]. Results ranged from incremental cost-effectiveness ratios of £203,900 to descriptions of increased cost of care [33]. With 17% (5/29) of studies referencing no improvement or reduced patient outcomes, the literature suggests that caregivers hesitate before providing telemonitoring care that is not cost-effective. Other barriers to consider are usability of devices, perceived lower autonomy of patients, and time required to obtain symptom data.

A total of 3 articles out of 29 (10%) explained prediction methods to determine the onset of exacerbations [9,19,25]. One article by Sanchez-Morillo et al [9] predicted 93.9% of exacerbations within 4.5 days of necessary medical interventions; they showed that improved patient outcomes and reduced in-patient visits can be achieved even with high variability of exacerbations. This article illustrates the variety of telemonitoring interventions available and an ideal method of protecting patient health through telehealth services.

Discussion

Principal Findings

A total of 12 facilitators [8,9,16,17,22,24-32,38,40] and 13 barriers [8,9,16-25,28,31-42] were identified in the literature and a total of 113 occurrences were detected. While frequency

does not impute importance, it does identify those issues most salient to those authors during the 6 years within which studies were published. Multiple factors were identified as both facilitators and barriers, but not by the same authors. The results of this review do not conflict with findings of other reviews, but they do imply several issues for consideration in health policy. Cost continues to play a reduced role in the use of telemedicine, which is a positive trend [33-35,38,39]. Policy makers should continue current incentives, but realize this may affect fewer organizations as the cost of implementation has already been absorbed. Lack of standardization is a barrier of concern and this issue is being addressed through organizations making and developing standards [18,21,22,24,28]. The most concerning and most frequently mentioned barrier was reduced patient outcomes or no improvement [16-18,20,28,31,38-42]. Technology is already expensive and it is often more complex than traditional care; decision makers seldom choose to pursue an intervention with technology unless there are improved patient outcomes that offset the cost of the technology itself. Policy makers need to carefully endorse those technology-infused interventions that yield positive patient outcomes and recommend that developers work on the rest of the barriers until the threshold for improvement is crossed.

Limitations

Authors noted some minor limitations. The high variability between articles, patient samples, telemonitoring methods, and treatments may explain why 24% (7/29) of the articles found improved patient outcomes and 21% (6/29) found no improvement in outcomes. With such differences between studied effects, external validity of the literature may be compromised. Further, because researchers were responsible for determining which articles were included in the study, readers should be aware of selection bias. However, biases were controlled by utilizing multiple reviewers for each article who discussed inclusion or exclusion of articles, as discussed in the Methods section.

Conclusions

Authors determined that many conflicting barriers and facilitators exist to the adoption of telemonitoring for patients with COPD. Due to the high variability of patients monitored, service lines, types of technology, and severity of disease state, some studies do not relate well to others. Future research should emphasize the importance of standardizing the telemonitoring of COPD techniques and the ability of technology to predict exacerbations. Predictability of exacerbations, even with the large range of pre-exacerbation symptoms, will reduce in-person visits and provide patients with useful warnings about their conditions. Through continued evaluation of COPD efficacy, research may find a cost-effective and useful standard for monitoring COPD through telehealth interventions.

Authors' Contributions

CK provided structure, guidance, and critical analysis; he analyzed 40% of the articles and rewrote portions of the manuscript. BP provided project management of the research; he assigned tasks, analyzed 40% of the articles, and wrote part of the Methods and Results sections. HC analyzed 30% of the articles, wrote the Introduction section, and created the first version of [Figure 1](#).

KB analyzed 30% of the articles and wrote the Discussion section. MA served as the lead editor; she analyzed 30% of the articles and wrote most of the Results section.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2017 Dec 01. Chronic obstructive pulmonary disease (COPD) URL: <http://www.who.int/mediacentre/factsheets/fs315/en/> [accessed 2017-03-18] [WebCite Cache ID 6p3V1bRuJ]
2. Feemster LC, Au DH. Penalizing hospitals for chronic obstructive pulmonary disease readmissions. *Am J Respir Crit Care Med* 2014 Mar 15;189(6):634-639 [FREE Full text] [doi: [10.1164/rccm.201308-1541PP](https://doi.org/10.1164/rccm.201308-1541PP)] [Medline: [24460431](https://pubmed.ncbi.nlm.nih.gov/24460431/)]
3. Stark P. Congressional intent for the HITECH Act. *Am J Manag Care* 2010 Dec;16(12 Suppl HIT):SP24-SP28 [FREE Full text] [Medline: [21314216](https://pubmed.ncbi.nlm.nih.gov/21314216/)]
4. Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth 2009. Geneva, Switzerland: World Health Organization; 2010. URL: https://www.who.int/goe/publications/goe_telemedicine_2010.pdf [accessed 2019-01-23] [WebCite Cache ID 75e7UJB2v]
5. Spradley P. Telemedicine: The law is the limit. *Tulane J Technol Intellect Prop* 2011;14:307-334 [FREE Full text]
6. Craig J, Patterson V. Introduction to the practice of telemedicine. *J Telemed Telecare* 2005;11(1):3-9. [doi: [10.1177/1357633X0501100102](https://doi.org/10.1177/1357633X0501100102)] [Medline: [15829036](https://pubmed.ncbi.nlm.nih.gov/15829036/)]
7. Bashshur RL, Shannon GW, Smith BR, Alverson DC, Antoniotti N, Barsan WG, et al. The empirical foundations of telemedicine interventions for chronic disease management. *Telemed J E Health* 2014 Sep;20(9):769-800 [FREE Full text] [doi: [10.1089/tmj.2014.9981](https://doi.org/10.1089/tmj.2014.9981)] [Medline: [24968105](https://pubmed.ncbi.nlm.nih.gov/24968105/)]
8. Jensen MH, Cichosz SL, Hejlesen OK, Toft E, Nielsen C, Grann O, et al. Clinical impact of home telemonitoring on patients with chronic obstructive pulmonary disease. *Telemed J E Health* 2012 Nov;18(9):674-678. [doi: [10.1089/tmj.2012.0003](https://doi.org/10.1089/tmj.2012.0003)] [Medline: [23134067](https://pubmed.ncbi.nlm.nih.gov/23134067/)]
9. Sanchez-Morillo D, Fernandez-Granero MA, Jiménez AL. Detecting COPD exacerbations early using daily telemonitoring of symptoms and k-means clustering: A pilot study. *Med Biol Eng Comput* 2015 May;53(5):441-451. [doi: [10.1007/s11517-015-1252-4](https://doi.org/10.1007/s11517-015-1252-4)] [Medline: [25725628](https://pubmed.ncbi.nlm.nih.gov/25725628/)]
10. Shea B, Grimshaw J, Wells G, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007 Feb 15;7:10 [FREE Full text] [doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10)] [Medline: [17302989](https://pubmed.ncbi.nlm.nih.gov/17302989/)]
11. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015 Jan 01;4:1 [FREE Full text] [doi: [10.1186/2046-4053-4-1](https://doi.org/10.1186/2046-4053-4-1)] [Medline: [25554246](https://pubmed.ncbi.nlm.nih.gov/25554246/)]
12. Ammenwerth E, Gräber S, Herrmann G, Bürkle T, König J. Evaluation of health information systems-problems and challenges. *Int J Med Inform* 2003 Sep;71(2-3):125-135. [Medline: [14519405](https://pubmed.ncbi.nlm.nih.gov/14519405/)]
13. Kruse C, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: A systematic review. *J Med Internet Res* 2015 Feb 10;17(2):e44 [FREE Full text] [doi: [10.2196/jmir.3171](https://doi.org/10.2196/jmir.3171)] [Medline: [25669240](https://pubmed.ncbi.nlm.nih.gov/25669240/)]
14. Light RJ. Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychol Bull* 1971;76(5):365-377. [doi: [10.1037/h0031643](https://doi.org/10.1037/h0031643)]
15. McHugh M. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
16. Alrajab S, Smith T, Owens M, Arenó J, Caldito G. A home telemonitoring program reduced exacerbation and healthcare utilization rates in COPD patients with frequent exacerbations. *Telemed J E Health* 2012 Dec;18(10):772-776. [doi: [10.1089/tmj.2012.0005](https://doi.org/10.1089/tmj.2012.0005)] [Medline: [23082792](https://pubmed.ncbi.nlm.nih.gov/23082792/)]
17. Antoniadis NC, Rochford PD, Pretto JJ, Pierce RJ, Gogler J, Steinkrug J, et al. Pilot study of remote telemonitoring in COPD. *Telemed J E Health* 2012 Oct;18(8):634-640. [doi: [10.1089/tmj.2011.0231](https://doi.org/10.1089/tmj.2011.0231)] [Medline: [22957501](https://pubmed.ncbi.nlm.nih.gov/22957501/)]
18. Burton C, Pinnock H, McKinstry B. Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease. *J Telemed Telecare* 2015 Jan;21(1):29-36. [doi: [10.1177/1357633X14562733](https://doi.org/10.1177/1357633X14562733)] [Medline: [25475218](https://pubmed.ncbi.nlm.nih.gov/25475218/)]
19. Celler BG, Sparks RS. Home telemonitoring of vital signs: Technical challenges and future directions. *IEEE J Biomed Health Inform* 2015 Jan;19(1):82-91. [doi: [10.1109/JBHI.2014.2351413](https://doi.org/10.1109/JBHI.2014.2351413)] [Medline: [25163076](https://pubmed.ncbi.nlm.nih.gov/25163076/)]
20. Chatwin M, Hawkins G, Panicchia L, Woods A, Hanak A, Lucas R, et al. Randomised crossover trial of telemonitoring in chronic respiratory patients (TeleCRAFT trial). *Thorax* 2016 Apr;71(4):305-311 [FREE Full text] [doi: [10.1136/thoraxjnl-2015-207045](https://doi.org/10.1136/thoraxjnl-2015-207045)] [Medline: [26962013](https://pubmed.ncbi.nlm.nih.gov/26962013/)]
21. Davis C, Bender M, Smith T, Broad J. Feasibility and acute care utilization outcomes of a post-acute transitional telemonitoring program for underserved chronic disease patients. *Telemed J E Health* 2015 Sep;21(9):705-713. [doi: [10.1089/tmj.2014.0181](https://doi.org/10.1089/tmj.2014.0181)] [Medline: [25955129](https://pubmed.ncbi.nlm.nih.gov/25955129/)]

22. Elwyn G, Hardisty AR, Peirce SC, May C, Evans R, Robinson DKR, et al. Detecting deterioration in patients with chronic disease using telemonitoring: Navigating the 'trough of disillusionment'. *J Eval Clin Pract* 2012 Aug;18(4):896-903. [doi: [10.1111/j.1365-2753.2011.01701.x](https://doi.org/10.1111/j.1365-2753.2011.01701.x)] [Medline: [21848942](https://pubmed.ncbi.nlm.nih.gov/21848942/)]
23. Fairbrother P, Pinnock H, Hanley J, McCloughan L, Sheikh A, Pagliari C, TELESCOT programme team. Continuity, but at what cost? The impact of telemonitoring COPD on continuities of care: A qualitative study. *Prim Care Respir J* 2012 Sep;21(3):322-328 [FREE Full text] [doi: [10.4104/pcrj.2012.00068](https://doi.org/10.4104/pcrj.2012.00068)] [Medline: [22875143](https://pubmed.ncbi.nlm.nih.gov/22875143/)]
24. Fairbrother P, Pinnock H, Hanley J, McCloughan L, Sheikh A, Pagliari C, et al. Exploring telemonitoring and self-management by patients with chronic obstructive pulmonary disease: A qualitative study embedded in a randomized controlled trial. *Patient Educ Couns* 2013 Dec;93(3):403-410. [doi: [10.1016/j.pec.2013.04.003](https://doi.org/10.1016/j.pec.2013.04.003)] [Medline: [23647981](https://pubmed.ncbi.nlm.nih.gov/23647981/)]
25. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. *Sensors (Basel)* 2015 Oct 23;15(10):26978-26996 [FREE Full text] [doi: [10.3390/s151026978](https://doi.org/10.3390/s151026978)] [Medline: [26512667](https://pubmed.ncbi.nlm.nih.gov/26512667/)]
26. Fernández-Granero MA, Sánchez-Morillo D, León-Jiménez A, Crespo LF. Automatic prediction of chronic obstructive pulmonary disease exacerbations through home telemonitoring of symptoms. *Biomed Mater Eng* 2014;24(6):3825-3832. [doi: [10.3233/BME-141212](https://doi.org/10.3233/BME-141212)] [Medline: [25227099](https://pubmed.ncbi.nlm.nih.gov/25227099/)]
27. Goldstein RS, O'Hoski S. Telemedicine in COPD: Time to pause. *Chest* 2014 May;145(5):945-949. [doi: [10.1378/chest.13-1656](https://doi.org/10.1378/chest.13-1656)] [Medline: [24798834](https://pubmed.ncbi.nlm.nih.gov/24798834/)]
28. Ho TW, Huang CT, Chiu HC, Ruan SY, Tsai YJ, Yu CJ, HINT Study Group. Effectiveness of telemonitoring in patients with chronic obstructive pulmonary disease in Taiwan: A randomized controlled trial. *Sci Rep* 2016 Mar 31;6:23797 [FREE Full text] [doi: [10.1038/srep23797](https://doi.org/10.1038/srep23797)] [Medline: [27029815](https://pubmed.ncbi.nlm.nih.gov/27029815/)]
29. Jordan R, Adab P, Jolly K. Telemonitoring for patients with COPD. *BMJ* 2013 Oct 17;347:f5932. [doi: [10.1136/bmj.f5932](https://doi.org/10.1136/bmj.f5932)] [Medline: [24136632](https://pubmed.ncbi.nlm.nih.gov/24136632/)]
30. Kim J, Kim S, Kim H, Kim K, Yang S, Lee C, et al. Effects of consumer-centered uHealth service for the knowledge, skill, and attitude of the patients with chronic obstructive pulmonary disease. *Comput Inform Nurs* 2012 Dec;30(12):661-671. [doi: [10.1097/NXN.0b013e318261c1c](https://doi.org/10.1097/NXN.0b013e318261c1c)] [Medline: [23266537](https://pubmed.ncbi.nlm.nih.gov/23266537/)]
31. Kim J, Kim S, Kim H, Kim K, Lee C, Yang S, et al. Acceptability of the consumer-centric uHealth services for patients with chronic obstructive pulmonary disease. *Telemed J E Health* 2012 Jun;18(5):329-338. [doi: [10.1089/tmj.2011.0140](https://doi.org/10.1089/tmj.2011.0140)] [Medline: [22510084](https://pubmed.ncbi.nlm.nih.gov/22510084/)]
32. Martín-Lesende I, Orruño E, Bilbao A, Vergara I, Cairo MC, Bayón JC, et al. Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on healthcare resource use (the TELBIL study randomised controlled trial). *BMC Health Serv Res* 2013 Mar 28;13:118 [FREE Full text] [doi: [10.1186/1472-6963-13-118](https://doi.org/10.1186/1472-6963-13-118)] [Medline: [23537332](https://pubmed.ncbi.nlm.nih.gov/23537332/)]
33. McDowell JE, McClean S, FitzGibbon F, Tate S. A randomised clinical trial of the effectiveness of home-based health care with telemonitoring in patients with COPD. *J Telemed Telecare* 2015 Mar;21(2):80-87. [doi: [10.1177/1357633X14566575](https://doi.org/10.1177/1357633X14566575)] [Medline: [25586812](https://pubmed.ncbi.nlm.nih.gov/25586812/)]
34. McKinsty B. The use of remote monitoring technologies in managing chronic obstructive pulmonary disease. *QJM* 2013 Oct;106(10):883-885. [doi: [10.1093/qjmed/hct068](https://doi.org/10.1093/qjmed/hct068)] [Medline: [23564633](https://pubmed.ncbi.nlm.nih.gov/23564633/)]
35. Pedone C, Chiurco D, Scarlata S, Incalzi RA. Efficacy of multiparametric telemonitoring on respiratory outcomes in elderly people with COPD: A randomized controlled trial. *BMC Health Serv Res* 2013 Mar 06;13:82 [FREE Full text] [doi: [10.1186/1472-6963-13-82](https://doi.org/10.1186/1472-6963-13-82)] [Medline: [23497109](https://pubmed.ncbi.nlm.nih.gov/23497109/)]
36. Pedone C, Lelli D. Systematic review of telemonitoring in COPD: An update. *Pneumonol Alergol Pol* 2015;83(6):476-484 [FREE Full text] [doi: [10.5603/PiAP.2015.0077](https://doi.org/10.5603/PiAP.2015.0077)] [Medline: [26559801](https://pubmed.ncbi.nlm.nih.gov/26559801/)]
37. Pinnock H, Hanley J, McCloughan L, Todd A, Krishan A, Lewis S, et al. Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: Researcher blind, multicentre, randomised controlled trial. *BMJ* 2013 Oct 17;347:f6070 [FREE Full text] [doi: [10.1136/bmj.f6070](https://doi.org/10.1136/bmj.f6070)] [Medline: [24136634](https://pubmed.ncbi.nlm.nih.gov/24136634/)]
38. Reddel HK, Jenkins CR, Partridge MR. Self-management support and other alternatives to reduce the burden of asthma and chronic obstructive pulmonary disease. *Int J Tuberc Lung Dis* 2014 Dec;18(12):1396-1406. [doi: [10.5588/ijtld.14.0371](https://doi.org/10.5588/ijtld.14.0371)] [Medline: [25517803](https://pubmed.ncbi.nlm.nih.gov/25517803/)]
39. Stoddart A, van der Pol M, Pinnock H, Hanley J, McCloughan L, Todd A, et al. Telemonitoring for chronic obstructive pulmonary disease: A cost and cost-utility analysis of a randomised controlled trial. *J Telemed Telecare* 2015 Mar;21(2):108-118. [doi: [10.1177/1357633X14566574](https://doi.org/10.1177/1357633X14566574)] [Medline: [25586810](https://pubmed.ncbi.nlm.nih.gov/25586810/)]
40. Venter A, Burns R, Hefford M, Ehrenberg N. Results of a telehealth-enabled chronic care management service to support people with long-term conditions at home. *J Telemed Telecare* 2012 Apr;18(3):172-175. [doi: [10.1258/jtt.2012.SFT112](https://doi.org/10.1258/jtt.2012.SFT112)] [Medline: [22362838](https://pubmed.ncbi.nlm.nih.gov/22362838/)]
41. Vianello A, Fusello M, Gubian L, Rinaldo C, Dario C, Concas A, et al. Home telemonitoring for patients with acute exacerbation of chronic obstructive pulmonary disease: A randomized controlled trial. *BMC Pulm Med* 2016 Dec 22;16(1):157 [FREE Full text] [doi: [10.1186/s12890-016-0321-2](https://doi.org/10.1186/s12890-016-0321-2)] [Medline: [27876029](https://pubmed.ncbi.nlm.nih.gov/27876029/)]

42. Zanamoni P, Lien LA, Hjalmsen A, Wootton R. Long-term telerehabilitation of COPD patients in their homes: Interim results from a pilot study in Northern Norway. *J Telemed Telecare* 2013 Oct;19(7):425-429. [doi: [10.1177/1357633X13506514](https://doi.org/10.1177/1357633X13506514)] [Medline: [24218358](https://pubmed.ncbi.nlm.nih.gov/24218358/)]

Abbreviations

AMSTAR: Assessment for Multiple Systematic Reviews

CINAHL: Cumulative Index to Nursing and Allied Health Literature

COPD: chronic obstructive pulmonary disease

HITECH: Health Information Technology for Economic and Clinical Health

MeSH: Medical Subject Headings

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

WHO: World Health Organization

Edited by G Eysenbach; submitted 05.07.18; peer-reviewed by J Hanley, M Duplaga; comments to author 08.10.18; revised version received 11.10.18; accepted 12.10.18; published 20.03.19.

Please cite as:

Kruse C, Pesek B, Anderson M, Brennan K, Comfort H

Telemonitoring to Manage Chronic Obstructive Pulmonary Disease: Systematic Literature Review

JMIR Med Inform 2019;7(1):e11496

URL: <http://medinform.jmir.org/2019/1/e11496/>

doi: [10.2196/11496](https://doi.org/10.2196/11496)

PMID: [30892276](https://pubmed.ncbi.nlm.nih.gov/30892276/)

©Clemens Kruse, Brandon Pesek, Megan Anderson, Kacey Brennan, Hilary Comfort. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 20.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Improving Digital Hospital Transformation: Development of an Outcomes-Based Infrastructure Maturity Assessment Framework

Patricia AH Williams^{1*}, BSc (Hons), GradCert, MSc, PhD; Brendan Lovelock^{2*}, BSc, PhD; Tony Cabarrus^{2*}; Marlon Harvey^{3*}

¹Flinders Digital Health Research Centre, College of Science and Engineering, Flinders University, Adelaide, Australia

²Cisco Systems Australia, Melbourne, Victoria, Australia

³Cisco Systems, Toronto, ON, Canada

* all authors contributed equally

Corresponding Author:

Patricia AH Williams, BSc (Hons), GradCert, MSc, PhD

Flinders Digital Health Research Centre

College of Science and Engineering

Flinders University

Flinders at Tonsley

GPO Box 2100

Adelaide, SA 5001

Australia

Phone: 61 82012023

Email: trish.williams@flinders.edu.au

Abstract

Background: Digital transformation in health care is being driven by the need to improve quality, reduce costs, and enhance the patient experience of health care delivery. It does this through both the direct intervention of technology to create new diagnostic and treatment opportunities and also through the improved use of information to create more engaging and efficient care processes.

Objective: In a modern digital hospital, improved clinical and business processes are often driven through enhancing the information flows that support them. To understand an organization's ability to transform their information flows requires a clear understanding of the capabilities of an organization's information technology infrastructure. To date, hospital facilities have been challenged by the absence of uniform ways of describing this infrastructure that would enable them to benchmark where they are and create a vision of where they would like to be. While there is an industry assessment measure for electronic medical record (EMR) adoption using the Healthcare Information and Management Systems Society Analytics EMR Adoption Model, there is no equivalent for assessing the infrastructure and associated technology capabilities for digital hospitals. Our aim is to fill this gap, as hospital administrators and clinicians need to know how and why to invest in information infrastructure to support health information technology that benefits patient safety and care.

Methods: Based on an operational framework for the Capability Maturity Model, devised specifically for health care, we applied information use characteristics to define eight information systems maturity levels and associated technology infrastructure capabilities. These levels are mapped to user experiences to create a linkage between technology infrastructure and experience outcomes. Subsequently, specific technology capabilities are deconstructed to identify the technology features required to meet each maturity level.

Results: The resulting assessment framework clearly defines 164 individual capabilities across the five technology domains and eight maturity levels for hospital infrastructure. These level-dependent capabilities characterize the ability of the hospital's information infrastructure to support the business of digital hospitals including clinical and administrative requirements. Further, it allows the addition of a scoring calculation for each capability, domain, and the overall infrastructure, and it identifies critical requirements to meet each of the maturity levels.

Conclusions: This new Infrastructure Maturity Assessment framework will allow digital hospitals to assess the maturity of their infrastructure in terms of their digital transformation aligning to business outcomes and supporting the desired level of clinical and operational competency. It provides the ability to establish an international benchmark of hospital infrastructure performance, while identifying weaknesses in current infrastructure capability. Further, it provides a business case justification through increased

functionality and a roadmap for subsequent digital transformation while moving from one maturity level to the next. As such, this framework will encourage and guide information-driven, digital transformation in health care.

(*JMIR Med Inform* 2019;7(1):e12465) doi:[10.2196/12465](https://doi.org/10.2196/12465)

KEYWORDS

medical informatics; information infrastructures; digital hospitals; eHealth; implementation; capability maturity modelling; security

Introduction

Background

Digital transformation in health care is being driven by the need to improve quality and reduce costs in health care delivery, while taking advantage of the benefits that advances in emerging technology can provide to patient care, the patient experience, workforce performance, and value and efficiency in health care delivery [1,2]. The recognition that digital transformation is important to health care is reflected in increasing discussions on what the digital health care future looks like and how we are moving to virtualized care venues, smart monitoring, and new trials using the Internet of things and cloud services [2,3]. Several studies have indicated that specific clinical initiatives require the improved support of health information technology infrastructure [4,5]. However, it is not the digital transformation itself that is the goal, but the capabilities and advances that this transformation can realize. While the fee-for-service health care funding model has been predominant in many countries, particularly outside of public hospitals, there is a shift to value-based health care funding with payment based on patient health outcomes. Hence, this shift necessitates assessment of the infrastructure to support new technologies and how the infrastructure aligns to clinical and administrative business value. Indeed, digital transformation will be critical to the survival of health care organizations [6].

To take advantage of technologies that can improve information flow in the digital hospital environment and enhance processes to create a positive impact on both clinical and operational outcomes, an assessment of the capabilities of the existing environment is required. Only through this process can we create a map for improvement and possibilities. Currently, there is no language to describe hospital technology infrastructure and no accepted standard methodology to assess the state of the network and infrastructure in a hospital to support structured improvement aligned to business processes. While there is such an industry assessment measure for electronic medical record (EMR) adoption using the Healthcare Information and Management Systems Society (HIMSS) Analytics EMR Adoption Model (EMRAM) [7], there is no equivalent for assessing the infrastructure and associated technology capabilities for digital hospitals. An infrastructure assessment framework is important to fill this gap as hospital administrators and clinicians need to know how and why to invest to support information infrastructure for the future. Such a framework will encourage and guide information-driven, digital transformation in health care.

This paper defines what information infrastructure a digital hospital requires to mature and use technology to enable

effective information flow and support the clinical and patient experience, now and in the future. The articulation and development process leading to the infrastructure maturity model and the information characteristics required to achieve the maturity objectives are given in the following section. This includes the mapping of systems capabilities to meet clinical process and patient experience, followed by a description of how we used a health care-specific capability mapping tool to devise a practical and industry best-practice assessment and scoring tool. The results demonstrate how this capability assessment mapping can be used to assess the maturity of a digital hospital to meet and improve its capabilities and how such a tool can be used for future planning of digital hospital technology requirements aligned to business and clinical outcomes.

Defining Information Maturity in Health Care

This section describes the reasoning behind, and formulation of, the maturity levels used to define infrastructure maturity, as well as the dimensions of information needed to place “information” in a position of enabling and supporting health care processes and decision making. It details the technology and technology services required to achieve those information dimensions and how those capabilities are aggregated into a set of naturally evolving levels. This description includes the research process of deconstruction of the associated technology requirements into a framework to determine the maturity of digital hospital infrastructure.

Information Dimensions to Enable and Support the Health Care Process

“Enabling information” means information that intrinsically possesses qualities to ensure it contributes effectively to health care delivery and decision making. This refers to both administrative and clinical data. The characteristics of enabling information in health care are analogous with data quality and comprise similar elements. Indeed, “information quality plays an important role as a mediator between clinical information technology and health care quality” [8]. Further, in different contexts, data quality may mean different things. For instance, administrative data are used both within the hospital environment for planning and funding, and also by the managing jurisdictions for services planning and policy making. Routine hospital data used outside the organization require standardization and homogenization for comparison within and across jurisdictions [9]. Further, data of lower quality can significantly influence business processes as well as clinical care [10]. A range of characteristics define information features and qualities that support patient care and its associated workflow processes [11]; this is a multidimensional concept

[8]. Russ et al [11] categorize such data quality characteristics into four domains: trustworthy and reliable, effectively displayed, adaptable to work demand, and ubiquitous. Each of these domains further defines the data characteristics that support that domain. For instance, the trustworthy and reliable domain consists of the data characteristics complete, consistent, correct, current, and secure. These characteristics are closely aligned to data quality characteristics.

In the context of health care, there are numerous definitions of what core quality characteristics should consist of, including accuracy, completeness, currency, and consistency, yet the associated characteristics of reliability, availability, usability, relevancy, and secureness are also important, in addition to trust, usefulness, and redundancy [10,12-16]. All these characteristics are accepted quality metrics throughout the literature. This paper does not, however, present a discourse on the ontology of data quality characteristics but instead accepts the premise that such characteristics exist and form the basis for extrapolation of the use of data with such embedded characteristics. These embedded characteristics provide the foundation for enabling information in health care.

There is an overlap between many of the data qualities. Therefore, each quality cannot be considered in isolation as a discrete concept because qualities come together in the context of use rather than through definitional articulation. There are numerous qualities that can describe data and information; however, in the health care environment these are not orthogonal or mutually exclusive. The selection of the following six higher-level characteristics used in this research reflects the usability of such characteristics when applied to health care:

1. **Completeness:** the information contains all the context required for decision making
2. **Relevancy:** the information required for the task at hand
3. **Usability:** information delivered on/in an appropriate device/format
4. **Availability:** information is available where it is required and exists in accessible systems
5. **Reliability:** information is sufficiently complete, error free, and consistent in distributed settings

6. **Security:** the information process is protected against extraction or tampering

These six characteristics give an insight at a high level into the importance of information in clinical application and are consistent with research into the use of patient information in electronic form [11] and how this influences operational and clinical workflow and service delivery. Further, each of the six dimensions have multiple interdependencies (see Table 1). Table 1 is based on the work of del Pilar and Garcia-Ugalde [14] with correlation from Russ et al [11] and Lee et al quality assessment characteristics [17].

A systematic review by Lee et al [17] identified accuracy as an intrinsic information quality across all studies analyzed. Using accuracy as one characteristic to explain these interdependencies, accuracy appears as a component of several dimensions but not one of the primary data quality dimensions in Table 1, because while accurate information is needed, on its own it does not enable health care delivery. However, accuracy is a primary element of the dimension of relevancy. Relevance refers to the information being appropriate for the task at hand, and information cannot be appropriate without being suitably accurate.

The interdependence of dimensions, such as accuracy and completeness (another intrinsic element that applies to multiple dimensions), demonstrates that attempting to definitively separate the dimensions is problematic and not productive. It should be noted that the dimension of “security” is a special case that supports all other dimensions.

The data quality dimensions (ie, characteristics of usable information) described in Table 1 are applied in this research to structure a scale of information maturity levels in an organization. They provide a loose set of guiding principles on how a hospital improves its use of information as it digitally matures. The characteristics are linked to the hospitals’ supporting information technology and technology services through the Capability Maturity Modeling (CMM) process described in the following section.

Table 1. Data quality dependencies adapted from [14].

Data quality dimension	Interdependencies
Completeness	<ul style="list-style-type: none"> • Coverage, density, relevancy, and sufficiency
Relevancy	<ul style="list-style-type: none"> • Current, timely, correct, and sufficient
Usability	<ul style="list-style-type: none"> • Usefulness consisting of relevance accuracy and completeness • Easy to use and organized
Availability	<ul style="list-style-type: none"> • Accessible, compatible, interpretable, and locatable
Reliability	<ul style="list-style-type: none"> • Unbiased • Reputation traceability including data source and provenance • Data producer with previous experience and correction of mistakes • Credibility inclusive of accuracy and completeness • Consistency
Security	<ul style="list-style-type: none"> • Supports all other dimensions

Capability Maturity Model

In the 1980s, CMM was devised as a tool to assess internal and external improvement processes [18], enabling a transformation from chaotic and ad hoc process implementation to definitive and disciplined process execution. CMM has an established background in software engineering, as well as applications to user-centered design [19], education [20], and information systems planning [21], for example.

The CMM methodology is based on ongoing improvement of capability and is constructed using:

- **Maturity levels.** These provide a structured template for persistent improvement. They promote the ability to assess new practice implication and success.
- **Key process areas and their associated goals.** A key process area is a set of related activities that can achieve the stated goal for that key process area.
- **Common features.** These are categories of key process areas and reflect the effectiveness and repeatability of that key process area. The CMM common features are commitment to performing, ability to perform, activities performed, measurement and analysis, and verifying implementation.
- **Key practices.** The implementation and persistent achievement in a key process area are defined by the procedures, practices, and communications implemented called key practices [22].

The adaptation of this CMM methodology into an operational framework (Figure 1) that can assess security technology, process, and human contribution in a medical environment in a simple and straightforward manner was developed by Williams [22].

This adaptation enables us to assess capability, competency, and maturity as a development of function, through the construction of a matrix of capability against competency that defines the maturity level reached (Table 2). The operational CMM framework (Figure 1) has been successfully applied to practical cybersecurity assessment of primary health care in Australia [23]. Table 2 is an example of this application in Australian primary health care, using back-up activities to articulate the key practice of back-up within the key process area of business continuity. As demonstrated in Table 2, the levels (1-5) define increasing competency in specific, deconstructed back-up activities.

A comprehensive discussion on the development of this framework can be found in Williams [22]. The framework was adapted to the context of this research as it relates to digital hospital infrastructure (described in the Methods section).

Enabling Information Mapping Using Capability Maturity Modeling

The HIMSS EMRAM maturity model, which measures the degree to which hospitals have replaced paper-based processes with technology, omits the ability to understand the supporting information and communication infrastructure required to

achieve each level of maturity. The HIMSS EMRAM maturity levels were used as the starting point from which to devise the capabilities required to support digital technologies to deliver optimal operational experience and business value. Using the principles of CMM, an eight-level (rather than the traditional five-level CMM) information systems Infrastructure Maturity Model (IMM) was developed (Figure 2). This was developed using Cisco Systems Australia's extensive experience in designing and building hospital infrastructure, by articulating the definition of infrastructure capability to support key domains, clinical services, and the application stack required to meet models such as HIMSS EMRAM. This was further enhanced through a series of observational ethnographic studies of hospital emergency departments that investigated how hospitals link the use of information with information technology and services that support a hospital's information-driven processes. These studies looked at the way technology is integrated into the clinical process, from simple to complex, with reference to the HIMSS EMRAM structure and level of sophistication required at each layer [24].

The IMM provides a framework for determining the preparedness of a hospital facility to support existing and planned process rollouts of digital infrastructure by classifying the way hospitals manage their digital infrastructure and by reflecting the sophistication of the information processes used within the organization. Each of the maturity levels is characterized in terms of the experiences they generate for key stakeholders (Table 3). The expected stakeholder experience at each of the maturity levels incorporates the patient experience, clinical process and quality, and associated productivity outcomes at a business level. This was tested, refined, and socialized with hospital Chief Information Officers through Cisco contacts to ascertain its relevancy and applicability.

Subsequently, the technology features and services (systems capability) for each maturity level were identified from the stakeholder experiences. These features demonstrate how technology can facilitate increased sophistication in health care delivery and the expected experience, reflecting the integration of technology and services (Table 4).

Summary

To demonstrate the business value of the IMM, Figure 3 shows how the initial five levels increase information access. Once the five levels have been achieved, the value can be seen through process and business improvement.

It is important to make the distinction between access value and process value. Improving the ability to access information enables the potential to create new and improved processes. However, these processes may not be realized because the supporting infrastructure that enables the application of this information may not exist. Articulation of the IMM into the Infrastructure Maturity Assessment (IMA) defines the outcomes a hospital could aspire to (ie, Levels 6-8) from enabling full utilization of the information accessed in Levels 1-5.

Figure 1. Representation of Capability Maturity Model as an operational framework.

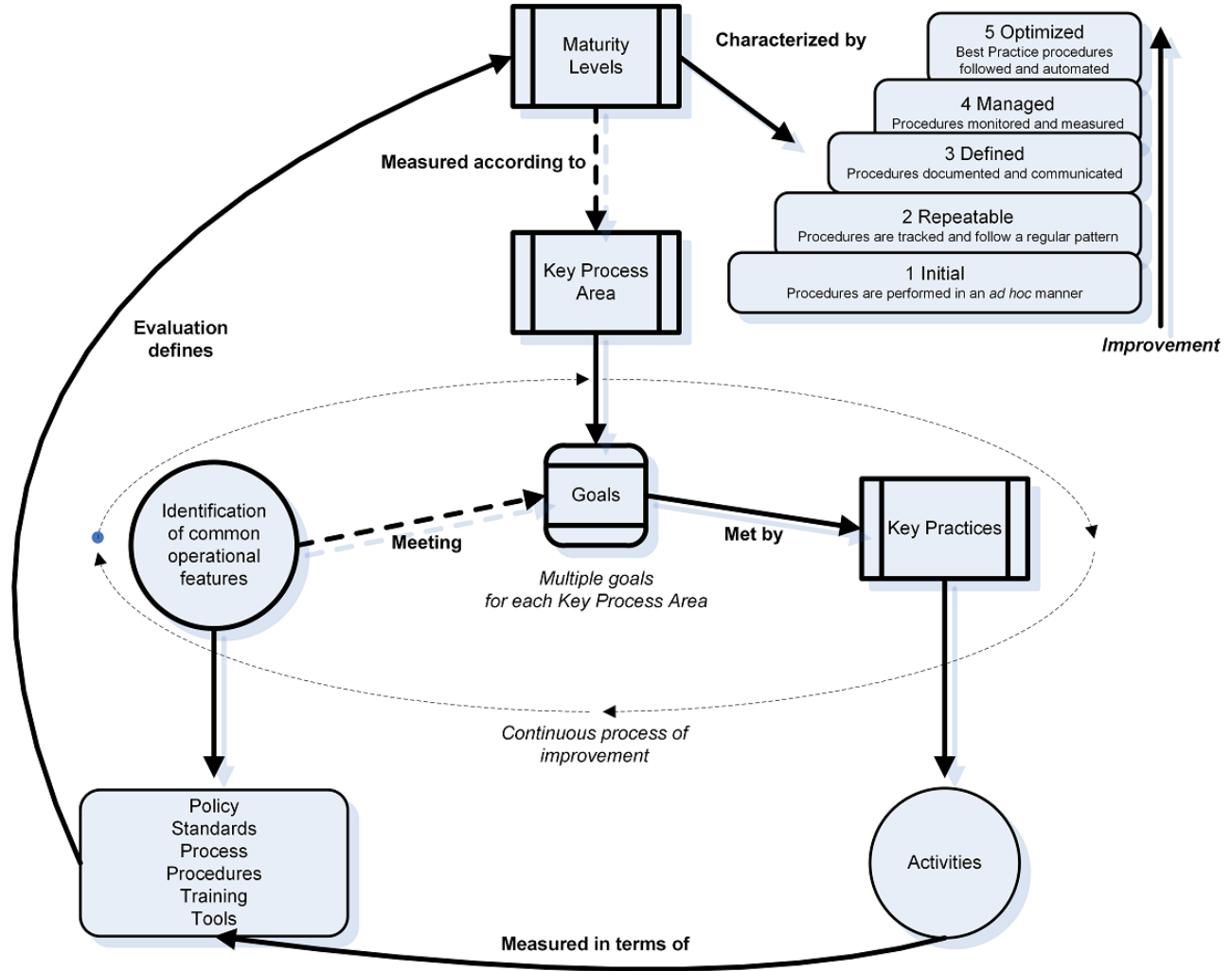


Table 2. Extract of operational Capability Maturity Modeling (CMM) matrix for back-up capability [22].

Back-up capability (activities)	Level 1 Initial	Level 2 Repeatable	Level 3 Defined	Level 4 Managed	Level 5 Optimized
Back-up frequency	None or manual initiation on ad hoc basis, or unknown	Manual initiation ad hoc, weekly, or every few days	Manual initiation daily	Automatic initiation daily	Automatic initiation. Continuous/real time with checks in place
Back-up type	None or partial (data only) or incremental	Partial (data and set-up files)	Full – all data	Full – all data and programs	Full systems back-up or imaging, including operating system
Back-up encryption	None	None	Encrypted	Encrypted with password	All back-ups encrypted and password-protected. Appropriate password protection control
Back-up reliability	None or back-up not checked, or unknown	Back-up checked for completion	Back-up periodically checked for reliability	Back-up periodically checked for reliability and outcome tracked	Back-up reliability tested with automatic notification. Every back-up outcome tracked

Figure 2. Information systems Infrastructure Maturity Model.

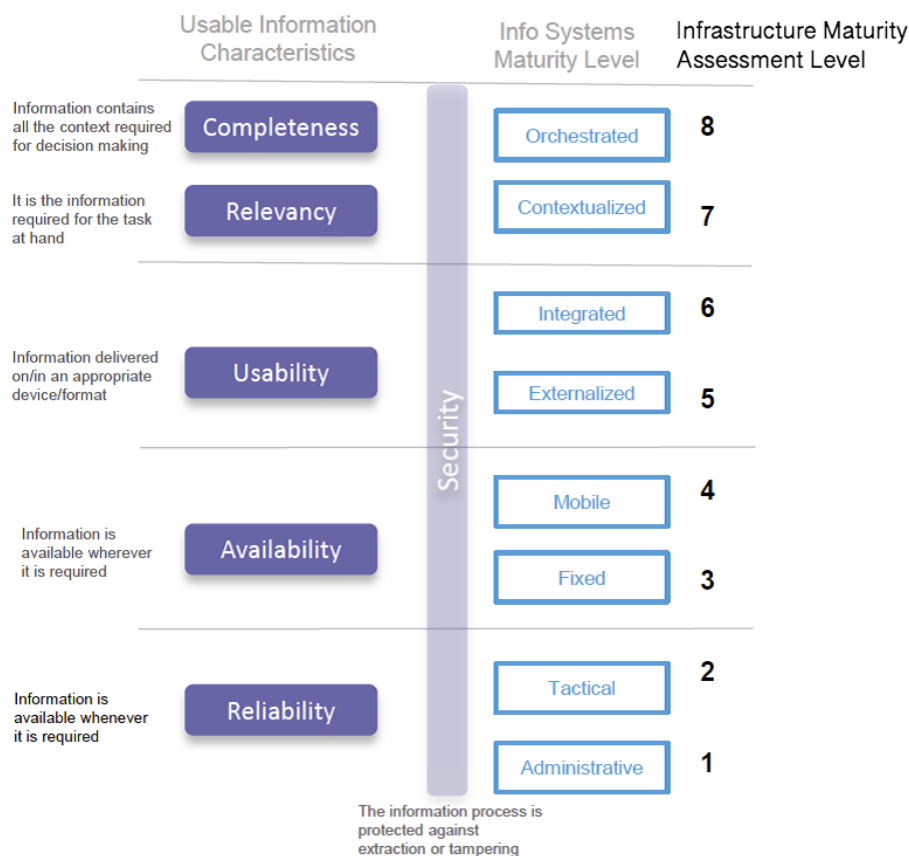


Table 3. Extract of stakeholder experience for the eight levels of data use maturity.

Maturity level for data use	Stakeholder experience description
Level 8: Orchestrated	The clinical and patient experiences can be molded not only to the role of the person but to their location, who is around them, and the requirements of the individual clinician or patient. The patient can be dynamically guided to where their next appointment is, advised if the appointment is running late, and prompted just before the doctor is ready to see them. They can be delivered educational material at the most appropriate time as well as advice on support services they may need as they exit the hospital. These types of services can come to their bedside terminal if they are a patient or to their personal phone if they are an inpatient. The same types of customized services can be delivered to clinical and operational staff in the hospital, enabling them to better manage their tasks and access the most important information or people they require for the task at hand.
Level 7: Contextualized	The clinical information is now customized to specific roles. There is a high level of data interoperability between clinical systems, and clinicians can get a single pane view of the patient. Task management and alerts are available and implemented according to operational and model of care requirements. Task management and alerts are closed loop, that is, there are escalation paths when tasks and alerts are not appropriately processed. Tasks and alerts are sent directly to the required individual’s mobile device rather than to their desktop. Patients can access information at their bedside terminal, which is customized to the individual patient’s needs. This includes building services such as catering, lighting, temperature, and other support services. Patient and staff needs can be centrally monitored and support delivered as required either from the nursing station or a centralized service delivery hub.
Level 2: Tactical	The hospital is starting to use information technology for clinical purposes. They have several clinical applications that are not linked (typically patient administration system [PAS], pharmacy, pathology, and radiology), and the network has sufficient speed to support these applications where they are required. There is a recognition of the importance of their PAS, and there are robust disaster recovery processes in place. The clinical applications are not always available to the clinical staff. Ordering results and general reporting are via paper and forms. The PAS system provides the central information resource. The information from the PAS is limited to a restricted number of operational and clinical staff. The requirements of the biometric devices in the facility have driven the deployment of data grade wireless where it is clinically required. The voice communications process is seen as an increasingly important element of clinical collaboration, and there is basic Internet Protocol telephony with a full featured console.
Level 1: Administrative	Hospitals do not use information technology for clinical use in any significant fashion. They do use information technologies for operational and financial purposes. These hospitals are paper-based in their clinical processes. They use fax, mail, and desk phones for communication and collaboration. Ordering and reporting are via forms. Information retrieval is via paper patient notes and internal paper courier services.

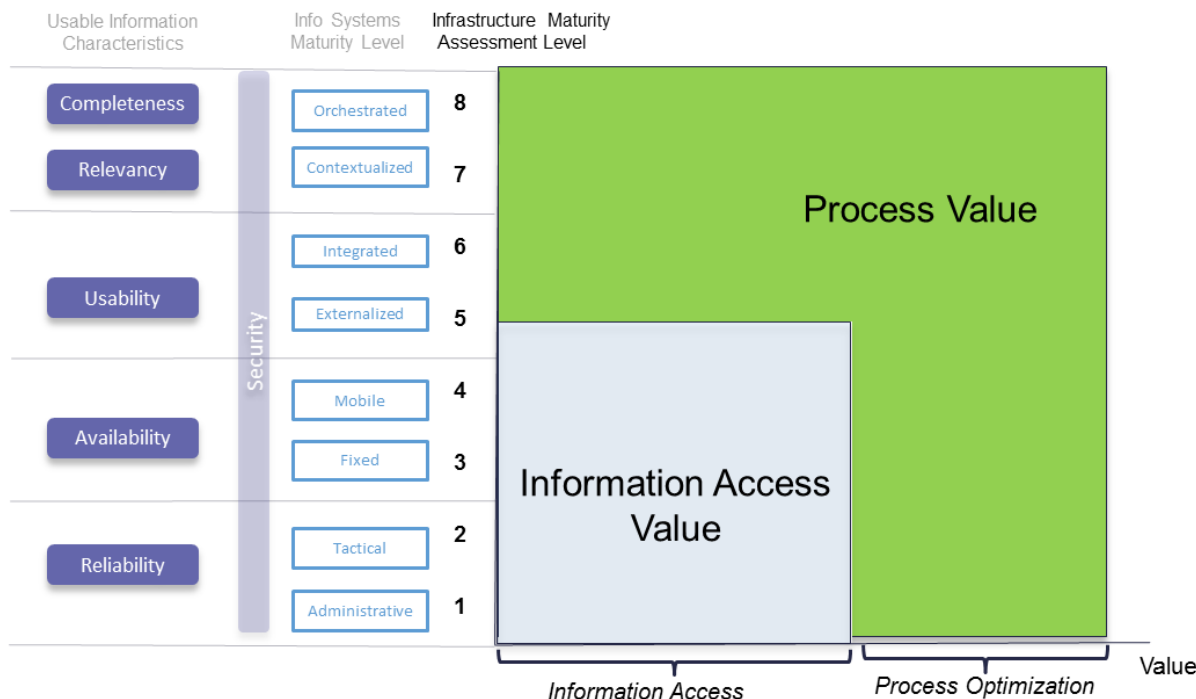
Table 4. Technology features and services in the Infrastructure Maturity Model (IMM).

Maturity level	Data use	Technology services
Level 8	Orchestrated	<ul style="list-style-type: none"> • Ability to link and coordinate processes in a centralized and automated fashion • Agile infrastructure, adaptable to the changing needs of the facility
Level 7	Contextualized	<ul style="list-style-type: none"> • Clinical processes customized to role and context • Closed loop alerts and tasks • Patient, staff, physical devices, and other resource location identification • Analytics and dynamic resource management
Level 6	Integrated	<ul style="list-style-type: none"> • Clinical processes on mobile devices • Combined info views for staff and patients • Bring your own device for staff and patients • Building Management Systems integration • Location services for key staff
Level 5	Externalized	<ul style="list-style-type: none"> • Ability to virtualize the major clinical and operational hospital services for delivery independent of location
Level 4	Mobile	<ul style="list-style-type: none"> • Clinical data available on mobile devices • Widely used mobile voice communications • Video services where needed • High level of collaboration services • Intelligent patient services • Duress services widely available • Locations services for equipment
Level 3	Fixed	<ul style="list-style-type: none"> • Broad digital clinical data availability • Ordering and reporting largely paper • Results online, clinical data repository • Integrated and distributed telephony services • High performance personal computers
Level 2	Tactical	<ul style="list-style-type: none"> • Department level apps to selected staff • Ordering/reporting/accessing are paper-based • Centralized high-quality telephony services
Level 1	Administrative	<ul style="list-style-type: none"> • Limited clinical applications • Paper-based systems • Analogue voice communications

The characterization of a hospital's infrastructure maturity using such a framework enables (1) identification of weaknesses in Information and Communications Technology (ICT) infrastructure capability, (2) definition of a business case justification of ICT investment, (3) provision of a roadmap for digital transformation in health care, and (4) measurement of international benchmarking of hospital infrastructure performance.

The usefulness of such a structure for characterizing health care ICT comes from the ability to link maturity levels and experience statements with the technology and services that need to be in place to deliver them. This research addresses the gap in the articulation of this technology infrastructure and services into practical and measurable capabilities.

Figure 3. Value outcomes of infrastructure maturity.



Methods

In order to translate the requirements into assessable capabilities, we used the Williams Operational CMM Framework (Figure 1), by contextualizing it for this research (Figure 4). The adaptation consisted of extending the number of maturity levels from 5 to 8, numbering the boxes and paths to reflect process steps below, and adding scoring components (6 Criticality and 7 Weighting). We also added text for correlated infrastructure terminology: Key Process Areas became (box 2) Key Process Areas (technology domains), Goals became (box 3) Goals (subdomains), Key Practices became (box 4) Key Practices (capabilities), and Activities became (box 5) Activities (assessable outcomes).

Using this framework (Figure 4), the process steps to create an operational infrastructure maturity assessment tool were as follows:

- Step 1 Maturity Levels: Define maturity levels, which means identifying and defining maturity levels appropriate to the target context (as described in Table 3).
- Step 2 Key Process Areas (Technology Domains): Identify key process areas (technology domains) means identifying the technology domains needed to support the outcomes for each of the maturity levels (from Step 1).
- Step 3 Goals (Subdomain Functions): Generate goals (subdomain function) comprises deconstructing each domain (from Step 2) into distinct, although not necessarily discrete functions, of that domain, called subdomains.

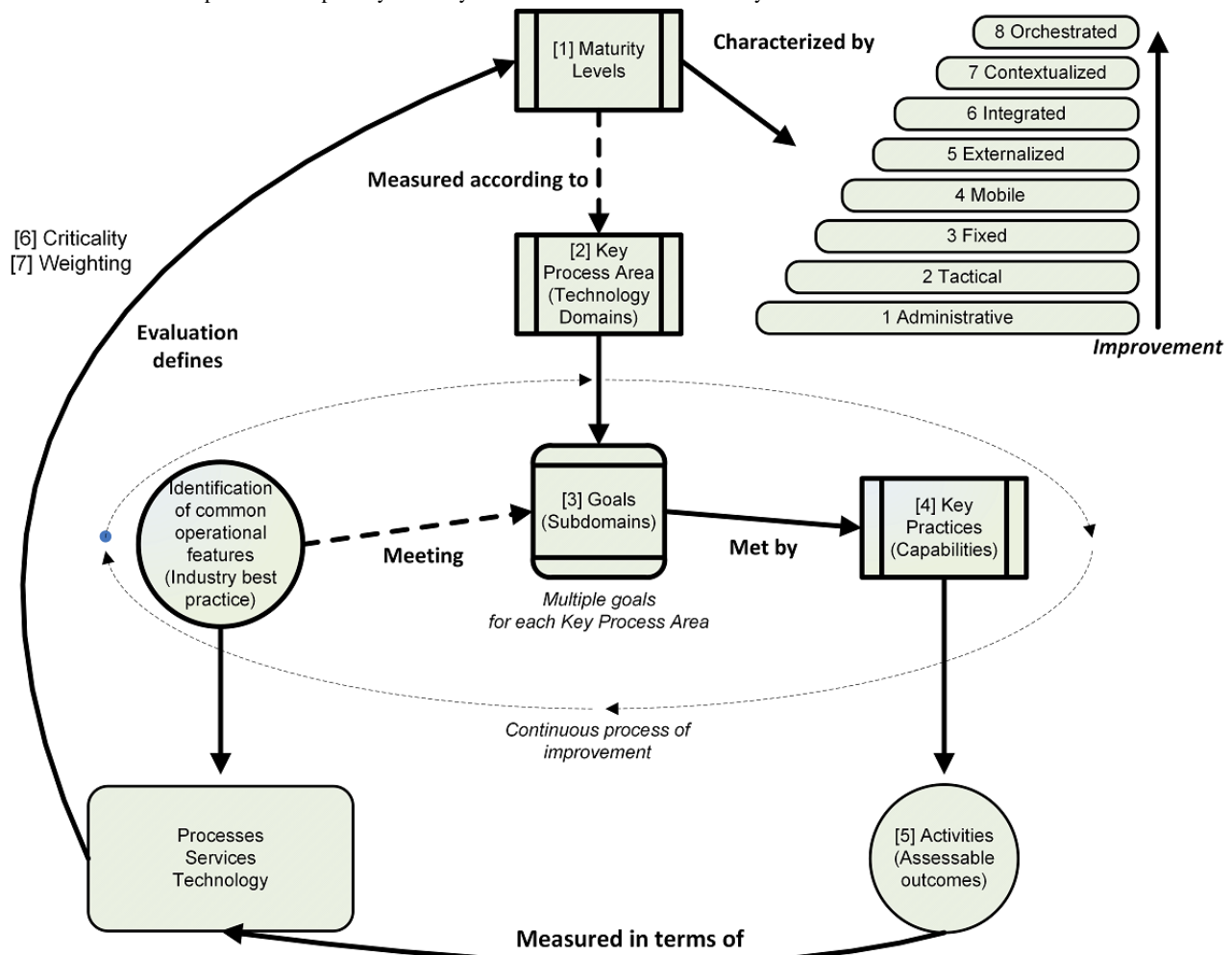
- Step 4 Key Practices (Capabilities): Devise key practices (capabilities) involves deconstructing each subdomain (from Step 3) into capabilities expected to meet each goal (subdomain function).
- Step 5 Activities (Measurable Outcomes): Articulate activities (measurable outcomes) means articulating measurable outcomes for each key practice capability (from Step 4) for each maturity level. The outcomes of this step also specify improvement from one maturity level to the next. This is represented as a capability matrix.

To facilitate a numeric scoring calculation of assessed infrastructure maturity, additional steps for assessment were devised.

- Step 6 Critical/Noncritical Capabilities: Identify the critical and noncritical capabilities for each key practice (from Step 4).
- Step 7 Assign Weightings: Assign weightings of importance to the goal subdomain (from Step 3) for each capability (from Step 4).

The research team consisted of the researcher, two Cisco infrastructure engineers and a health care technology expert, all with extensive experience in health care and the hospital environment in Australia and the United States. The construction of the framework was purposely constructed as nonproprietary and therefore industry generalizable. All capabilities researched and found to be Cisco specific were omitted from the capabilities. These capabilities may be explored for future revisions of the framework and for driving future technical innovation.

Figure 4. Contextualized operational Capability Maturity Model for infrastructure maturity assessment.



Results

Step 1: Maturity Levels

Eight maturity levels were identified and defined as appropriate to the digital hospital context. This definition and underlying reasoning for the eight levels of infrastructure maturity are explained in Table 3.

Step 2: Key Process Areas (Technology Domains)

In defining the technology landscape to articulate the key process areas (technology domains), existing categories of outcome-based functionality were apparent. To manage the large scale of a hospital landscape, it was logical to group the technology and services into existing categories of outcomes-based functionality. These categories reflect five technology domains: (1) Transport, (2) Collaboration, (3) Security, (4) Mobility, and (5) Data Center.

Step 3: Goals (Subdomain Functions)

The subdomains are the division of each domain into distinct, although not necessarily discrete (for that domain), functions of that domain. The methodology, by definition, looks at the

delivery of services through infrastructure and therefore includes some services as infrastructure.

To illustrate the results of this methodology, the subdomain, “Transport” is used as an example. In the Transport domain, six specific functions (subdomains) of transport in infrastructure were identified: (1) Campus Connectivity, (2) Secure Remote Access, (3) Traffic Optimization–Quality of Service (QoS), (4) Disruption Tolerance and High Availability, (5) Management, and (6) Extensibility.

Step 4: Key Practices (Capabilities)

The technology subdomains are further subdivided into key practices called capabilities. Each capability is described by a set of related information systems outcomes. The technology and technology services required to deliver those outcomes are then sequenced from Levels 1-8 of the IMA framework. Table 5 is an extract from the Transport domain using the Campus Connectivity subdomain and three of the capabilities from this subdomain together with the measures of capability (activities) at each maturity level. Levels 1 and 8 are presented as examples, together with the abbreviated descriptor of each capability.

Table 5. Extract of Transport domain, Campus Connectivity subdomain, with capability descriptors and measurable outcomes.

Capability	Descriptor (abbreviated)	Level 1 ^a	Level 8
Cabling standard	ANSI/TIA-1179-A “Healthcare Facility Telecommunications Infrastructure” specifies recognized cabling and cabling category recommendations for health care facilities.	Category 6A cabling is ≤30%	Category 6A cabling >91%
Virtualization	The concept of virtualization applies tags to network packets that create the appearance and functionality of network traffic that is physically on a single network but acts as if it is split between separate networks.	No layer 2/3 virtualization implemented	Access controlled, policy-based micro segmentation of campus infrastructure based on virtual extensible local area network
Access port design and policy	A well-defined access port policy is based on the requirements of the end devices and the access of the applications and services by that end device.	No access port policy	Software defined automation access port configuration per software defined networking (SDN) policy
SDN integration	The SDN controller should support integration using application programming interfaces (APIs). Representational State Transfer (REST) APIs enable automation, integration, and innovation. All controller functionality should be exposed through these REST APIs.	No SDN	Demonstration of SDN contextual workflow using API integration

^aLevels 2-7 (no entries) define increasing maturity assessment criteria.

Step 5: Activities (Measurable Outcomes)

The activities (Table 5) express measurable outcomes (technology/technology services) for each capability at each maturity level, thus creating a capability matrix. In doing so, this matrix also specifies improvement from one maturity level to the next, which may be through increasing functionality, quality, or provision of a technology/service. The placement of each measurable outcome to a maturity level is based on best practice, in-depth knowledge, and experience by the Cisco design and implementation engineers and the external contributor to the project with expertise in health care technologies, networking, and CMM in the health care environment.

Step 6: Critical / Noncritical Capabilities

As demonstrated throughout this paper, a digital hospital’s infrastructure is complex with many facets. To provide a collated assessment of the maturity to meet the business goals, a measure of essential and nonessential capability is required. Once these capability measures were defined, the critical and noncritical capabilities were identified for each capability within a domain and subdomain. These decisions were based on the importance of a specific capability to meet the business outcomes at that maturity level and the risk to the associated service delivery in not achieving this capability at this level. The critical capabilities are defined as essential criteria and a requirement that hospital facilities must meet to be placed at this level, regardless of other components reached at that level. If a capability is noncritical, then it is desirable but not essential for the maturity level.

Step 7: Assign Weightings

Weightings are also assigned to each capability reflecting the importance of the capability within the subdomain. This is in addition to the critical and noncritical status of the capability because the criticality measure is a binary measure, and where there is more than one critical and noncritical capability, their importance to the subdomain is not necessarily equal. This allows for the more granular calculation of weighting

specifically related to the maturity levels and the expectations in a hospital environment of what that maturity level would allow the organization to undertake both clinically and administratively.

Summary

The collective set of matrices created using this methodology make up the IMA framework. Across the framework, there are a total of five domains, 34 subdomains, and 164 individual capabilities defined (see Multimedia Appendix 1).

Discussion

Principal Considerations

This research resulted in the identification of five key process area technology domains (Transport, Collaboration, Security, Mobility, and Data Center) each comprising multiple distinct subdomains. The 34 subdomains define functionality spread across the five technology domains. The articulation of this functionality into 164 measurable capabilities was achieved through the definition of each capability at each of the eight maturity levels.

We anticipate that the framework would be used to guide organizations as they go through major digital transformations. These transformations can be driven by the implementation of major clinical systems (such as EMRs), the refurbishment of brownfield facilities, or the building of new hospitals or area health services.

The assessment of capabilities is an uncomplicated exercise to match current practice to the best-fit maturity level, although technical knowledge of the implemented infrastructure environment is required. Each domain/subdomain identifies the capabilities that need to be in place. The measurable outcomes specify in detail how this is undertaken, reflecting the competency and maturity levels.

The challenges in defining and deconstructing each domain and its capability included a common understanding of terminology

from a technical perspective and expansion of this for a broader technology informed audience. In developing a capabilities matrix, particularly one with the necessary eight levels, the challenge is that some capabilities may not change across each individual level. Therefore, assigning a level for that capability and subsequently calculating its weighting can be problematic. Where this occurs in the assessment, the highest level that capability can be assigned is allocated. This does not impact the weighting calculation for the level attained because the highest level is fundamentally determined by whether the critical capabilities for the domain or subdomain are met.

The process of using the IMA framework consists of (1) analysis of the hospital's information systems infrastructure across the five technology domains, (2) classification matched to the eight-level maturity model against the relevant operational outcomes, (3) a roadmap and tailored objectives for each technology area outlining efforts needed to improve capability, (4) comparison and benchmarking of a digital hospital's information infrastructure capabilities, and (5) recommendations for achieving business and IT goals to meet business and experience outcomes.

Infrastructure Maturity Assessment Outputs

The following examples demonstrate typical outputs of the IMA framework and how they could be used by C-Suite executives (Chief Executive Officer, Chief Information Officer, Chief Marketing Officer, Chief Medical Informatics Officer) as a reference point for decision making on infrastructure investment.

The IMA output demonstrates the ability of the framework to create a benchmark for measuring the capabilities of a given hospital or health service. Further breaking the result down into technology subdomains and capabilities provides a detailed fingerprint of the ICT capabilities and how they relate to the requirements and aspirations of the health care provider.

Figure 5 provides a picture of the current state of maturity across 25 de-identified Australian hospitals and demonstrates that most hospitals are operating at Level 3 (Fixed) infrastructure maturity. This level of capability creates significant limitations in the ability of the hospital facility to take advantage of technologies aimed at improving information flow and process operations within the hospital facility through achieving Level 4 (Mobile).

Level 3 describes an organization with limitations in its wireless, transport, and collaboration capability, restricting the ability to reliably access, share, and act on clinical information throughout the facility. These limitations diminish the value of the information applications that the organization has invested in by restricting the availability of information to when the clinical staff have access to personal computer endpoints. Ultimately, this limits the ability to drive clinical and operational process automation that can be gained from making data mobile.

As an example, Figure 6 demonstrates a maturity assessment result for the Transport domain of one hospital, showing the individual result for each of the subdomains in the Transport domain.

The potential improvement lies in incrementally improving each subdomain infrastructure capability (as defined in Table 5), thus increasing the value proposition to the next level. Analysis of each subdomain score allows reflection on the current infrastructure capability with the typically experienced limitations highlighted as follows using the Transport example:

- The impact of the subdomain assessment scores needs careful consideration when assessing the value proposition of infrastructure investment. For instance, the high Campus Connectivity score indicates a Transport infrastructure that is approaching high performance capability. However, the lower scores in the associated subdomains indicate potential network unreliability, suboptimal management of data priority, and network failure. This defines an infrastructure that is unable to fully leverage its capabilities.
- While the challenges in Campus Connectivity faced by many hospital facilities are less reflected in this assessment result, challenges commonly include aged infrastructure particularly at the Access layer (5+ years), legacy campus designs that are inflexible, separate physical networks for medical devices and building management systems, a lack of properly defined segmentation policies and design, a lack of network visibility for contextual behavioral analytics, and a lack of API integration capability. These challenges collectively result in a deficiency to support advanced functionality, modularity, and scalability, which is a key requirement to support data use at higher maturity levels.
- The challenges in Traffic Optimization–QoS (a maturity Level 2 in Figure 6) are primarily a result of fragmented QoS designs, the problematic processes to design and implement new campus and wide area network QoS configurations, and no clear roadmap to define an End-to-End contextual Quality of Experience. A lower Traffic Optimization score is often the result of an isolated QoS approach that does not cope well with new services and the associated traffic prioritization requirements.
- Similarly, in the Disruption Tolerant Networking & High Availability subdomain (scoring Level 1 in Figure 6), low scores are often a result of potential for single points of failure in the Core/ Distribution and Access layers of the infrastructure, where devices are locally configured, arising from a lack of comprehensive designs based on Application and Infrastructure interaction.

These challenges provide a snapshot of how the IMA framework can describe the current architecture, identify the misalignment between existing capability and desired capability, and inform subsequent infrastructure planning.

Figure 5. Example of the overall maturity assessment measure across 25 hospitals in Australia.

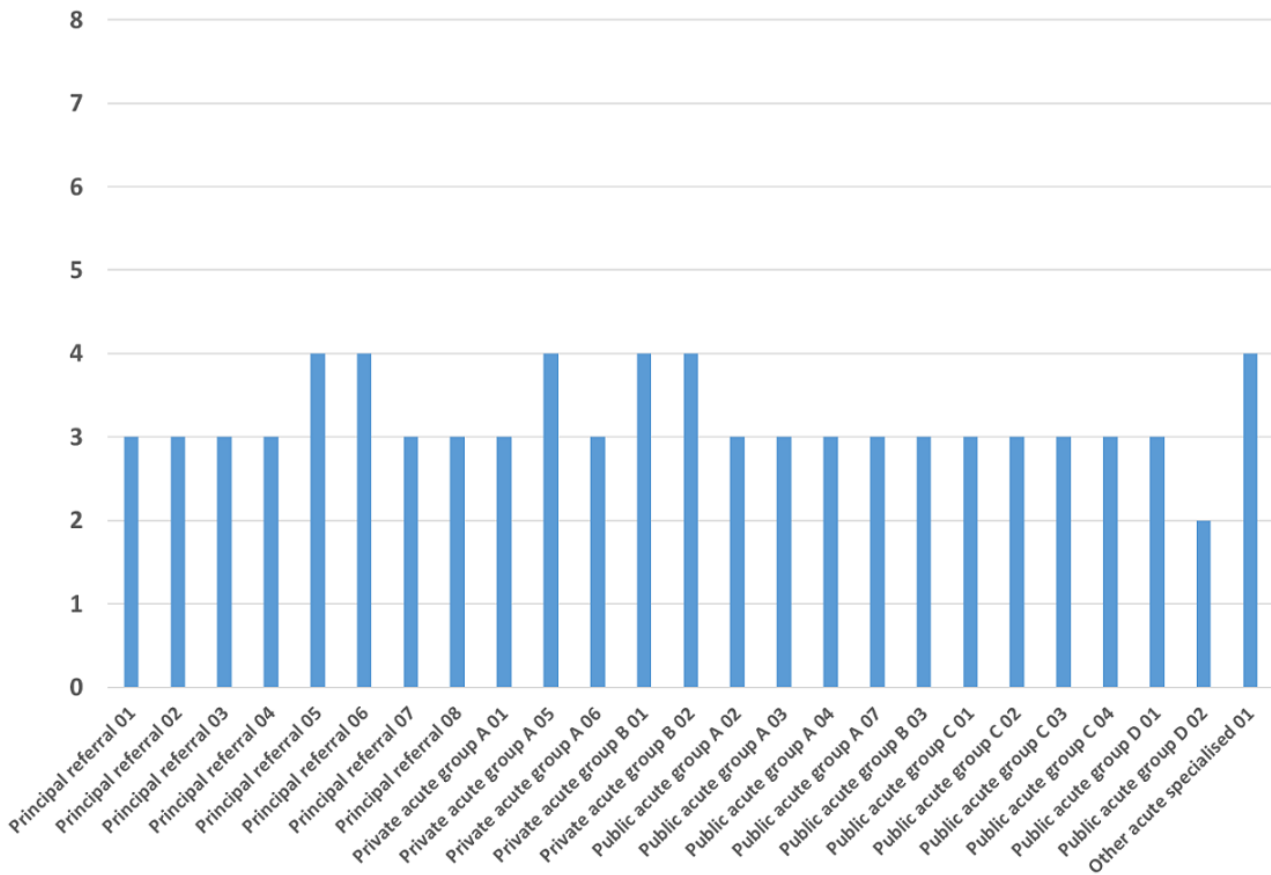
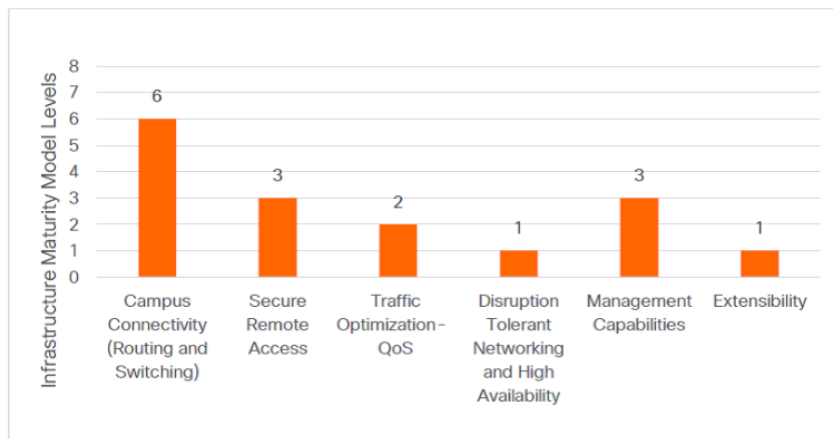


Figure 6. Example of Transport domain assessment output. QoS: quality of service.



Framework Extensions

One application of the IMA framework is that it could be mapped to the HIMSS EMRAM, as both frameworks provide a roadmap to support broader process changes in hospitals. Arguably the IMA framework levels could correspond to the application stack requirements to meet the EMRAM model. This would provide a new health care industry benchmark of the maturity of the clinical application deployment and use within hospitals that can be supported by a given level of ICT infrastructure maturity. Localization may be required in applying this framework to countries outside Australia.

Conclusion

Digital hospitals need to take advantage of the technologies that can improve information flow and use to meet quality clinical and administrative outcomes. The necessity for the technology infrastructure to support these outcomes is clear. Yet such infrastructure is complex and continually evolving in its design and deployment particularly when this involves many stakeholders’ demands and expectations. The resultant assessment of such infrastructure to meet business outcomes and realize value to health care organization through its capability is equally complex.

The use of the Williams CMM Operational Framework allows such capability to be deconstructed into manageable constituent elements and assessed individually. Through this process, it also identifies specific incremental improvement for each capability. The resulting IMA framework allows hospital management and technicians to clearly see how incremental improvements in their infrastructure can be achieved to support clinical and operation goals. Such a method assists hospitals to define an improvement pathway and maturity in delivering their organizational objectives.

The IMA characterizes the technology services required to support a hospital's information-driven processes. Thus, it provides a tool for determining the preparedness of a hospital facility to support existing or planned process rollouts. The IMA classifies the way hospitals manage their digital infrastructure into an eight-level model that reflects the sophistication of the information processes used within the facility. Each of these levels is characterized in terms of the experiences they generate for the key stakeholders and the technology services required to support those experiences.

Importantly, the IMA framework articulates how hospitals can generate more value from their infrastructure as it defines the levels at which the critical "enabling information" characteristics for an organization are primarily delivered. It describes the stages of information use and resultant clinical/patient experience within a hospital and the ICT infrastructure requirements to reliably achieve those levels of experience. The framework is designed to stage the infrastructure development pathway so that clear benefits can be attributed to the incremental investment that is required to progress from one stage to the next. Consequently, robust business cases can be made for that investment. Ultimately, the purpose of the framework is to map a pathway where Chief Information Officers can see a logical sequence of infrastructure development that they need to take their hospital facility through to reach their desired level of clinical and operational competency.

The development of a generalizable Infrastructure Maturity Assessment tool contributes to and supports the digital hospital industry, providing an international benchmarking standard.

Acknowledgments

Cisco Systems Australia is acknowledged as sponsor of this research.

Conflicts of Interest

BL and TC are employed by Cisco Systems Australia. MH is employed by Cisco Systems. PAHW is Cisco Chair at Flinders University and received a grant for this research.

Multimedia Appendix 1

Infrastructure Maturity Assessment framework capabilities.

[[PDF File \(Adobe PDF File\), 73KB - medinform_v7i1e12465_app1.pdf](#)]

References

1. Agarwal R, Gao G, DesRoches C, Jha AK. The Digital Transformation of Healthcare: Current Status and the Road Ahead. *Information Systems Research* 2010 Dec;21(4):796-809. [doi: [10.1287/isre.1100.0327](https://doi.org/10.1287/isre.1100.0327)]
2. SAP. 2018 May 4. A future in digital health 2017 URL: <https://www.sap.com/industries/healthcare.html> [accessed 2018-10-10] [[WebCite Cache ID 733iSNn1u](#)]
3. Ashton K. Making sense of IoT: How the Internet of Things became humanity's nervous system. 2017 Feb 10. URL: <https://www.scribd.com/document/346619901/HPE-Aruba-IoT-eBook> [accessed 2018-10-10] [[WebCite Cache ID 75J7R3ud3](#)]
4. Trtovac D, Lee J. The Use of Technology in Identifying Hospital Malnutrition: Scoping Review. *JMIR Med Inform* 2018 Jan 19;6(1):e4 [[FREE Full text](#)] [doi: [10.2196/medinform.7601](https://doi.org/10.2196/medinform.7601)] [Medline: [29351894](https://pubmed.ncbi.nlm.nih.gov/29351894/)]
5. Choi I, Kim JK, Kim SJ, Cho SC, Kim IN. Satisfaction Levels and Factors Influencing Satisfaction With Use of a Social App for Neonatal and Pediatric Patient Transfer Information Systems: A Questionnaire Study Among Doctors. *JMIR Med Inform* 2016 Aug 04;4(3):e26 [[FREE Full text](#)] [doi: [10.2196/medinform.5984](https://doi.org/10.2196/medinform.5984)] [Medline: [27492978](https://pubmed.ncbi.nlm.nih.gov/27492978/)]
6. Walsh M, Rumsfeld J. Leading the Digital Transformation of Healthcare: The ACC Innovation Strategy. *J Am Coll Cardiol* 2017 Nov 28;70(21):2719-2722 [[FREE Full text](#)] [doi: [10.1016/j.jacc.2017.10.020](https://doi.org/10.1016/j.jacc.2017.10.020)] [Medline: [29169479](https://pubmed.ncbi.nlm.nih.gov/29169479/)]
7. HIMSS Analytics. 2018. Electronic Medical Record Adoption Model: HIMSS Analytics EMRAM URL: <https://www.himssanalytics.org/emram> [accessed 2018-10-10] [[WebCite Cache ID 733iiwXJG](#)]
8. Byrd L, Byrd T. Developing an instrument for information quality for clinical decision making. : IEEE; 2012 Jan 04 Presented at: 45th Hawaii International Conference on System Sciences; Jan. 4-7, 2012; Maui, Hawaii URL: <https://ieeexplore.ieee.org/document/6149169> [doi: [10.1109/HICSS.2012.210](https://doi.org/10.1109/HICSS.2012.210)]
9. Bernal-Delgado E. Data quality assurance approaches using administrative data in healthcare performance assessment. *European Journal of Public Health* 2017;27(3):- [doi: [10.1093/eurpub/ckx187.755](https://doi.org/10.1093/eurpub/ckx187.755)]
10. Batini C, Scannapieco M. Data quality dimensions. In: Batini C, editor. *Data and Information Quality. Data-Centric Systems and Applications*. Cham, Switzerland: Springer International Publishing; 2016:21-51.

11. Russ A, Saleem J, Justice C, Woodward-Hagg H, Woodbridge P, Doebbeling B. Electronic health information in use: Characteristics that support employee workflow and patient care. *Health Informatics J* 2010 Dec;16(4):287-305. [doi: [10.1177/1460458210365981](https://doi.org/10.1177/1460458210365981)] [Medline: [21216808](https://pubmed.ncbi.nlm.nih.gov/21216808/)]
12. Batini C, Scannapieco M. Information Quality Dimensions for Maps and Texts. In: Batini C, Scannapieco M, editors. *Data and Information Quality. Data-Centric Systems and Applications*. Cham, Switzerland: Springer International Publishing; 2016:53-86.
13. Cabitza F, Batini C. Information quality in healthcare. In: Batini C, Scannapieco M, editors. *Data and Information Quality. Data-Centric Systems and Applications*. Cham, Switzerland: Springer International Publishing; 2016:403-419.
14. del Pilar Angeles M, Garcia-Ugalde FJ. Subjective assessment of data quality considering their interdependencies and relevance according to the type of information systems. *International Journal on Advances in Software* 2012;5(3-4):389-400 [[FREE Full text](#)]
15. Organisation for Economic Co-operation and Development. OECD Reviews of health care quality: Australia 2015 URL: <http://www.oecd.org/australia/oecd-reviews-of-health-care-quality-australia-2015-9789264233836-en.htm> [accessed 2018-10-10] [[WebCite Cache ID 75J8Krnxy](#)]
16. Sebastian-Coleman L. *Measuring Data Quality For Ongoing Improvement. A Data Quality Assessment Framework*. Burlington, MA: Morgan Kaufmann Publishers Inc; 2013.
17. Lee YW, Strong DM, Kahn BK, Wang RY. AIMQ: a methodology for information quality assessment. *Information & Management* 2002 Dec;40(2):133-146. [doi: [10.1016/s0378-7206\(02\)00043-5](https://doi.org/10.1016/s0378-7206(02)00043-5)]
18. Galin D, Avrahami M. Are CMM Program Investments Beneficial? Analyzing Past Studies. *IEEE Softw* 2006;23(6):81-87. [doi: [10.1109/MS.2006.149](https://doi.org/10.1109/MS.2006.149)]
19. Jokela T. Evaluating the user-centredness of development organisations: conclusions and implications from empirical usability capability maturity assessments. *Interacting with Computers* 2004 Dec;16(6):1095-1132. [doi: [10.1016/j.intcom.2004.07.006](https://doi.org/10.1016/j.intcom.2004.07.006)]
20. Boulet M, Dupuis C, Belkhter N. Selecting continuous training program and activities for computer professionals. *Computers & Education* 2001 Jan;36(1):83-94. [doi: [10.1016/S0360-1315\(00\)00057-9](https://doi.org/10.1016/S0360-1315(00)00057-9)]
21. Saleh Y, Alshawi M. An alternative model for measuring the success of IS projects: the GPIS model. *Journal of Ent Info Management* 2005 Feb;18(1):47-63. [doi: [10.1108/17410390510571484](https://doi.org/10.1108/17410390510571484)]
22. Williams P. A practical application of CMM to medical security capability. *Info Mngmnt & Comp Security* 2008 Mar 21;16(1):58-73. [doi: [10.1108/09685220810862751](https://doi.org/10.1108/09685220810862751)]
23. RACGP. *Computer and information security standards and workbook*. Melbourne, Australia: Royal Australian College of General Practitioners; 2012. URL: <https://www.racgp.org.au/FSDEDEV/media/documents/Running%20a%20practice/Practice%20standards/Computer-and-information-security.pdf> [accessed 2018-10-10] [[WebCite Cache ID 75J9C6cLO](#)]
24. Creative Intellect Consulting. *The network foundation for IOT and connected healthcare*. 2018. CIC Healthcare Series URL: <https://www.creativeintellectuk.com/download-from-cic/healthcare/progressing-the-patient-experience-in-iot-and-connected-healthcare/> [accessed 2018-10-09] [[WebCite Cache ID 733jiloEq](#)]

Abbreviations

- API:** application programming interface
- CMM:** Capability Maturity Model
- EMR:** electronic medical record
- EMRAM:** EMR Adoption Model
- HIMSS:** Healthcare Information and Management Systems Society
- ICT:** information and communications technology
- IMA:** Infrastructure Maturity Assessment
- IMM:** Infrastructure Maturity Model
- PAS:** patient administration system
- QoS:** quality of service
- REST:** Representational State Transfer
- SDN:** software defined networking

Edited by G Eysenbach; submitted 10.10.18; peer-reviewed by V McCauley, Y Qiu; comments to author 01.11.18; revised version received 29.11.18; accepted 09.12.18; published 11.01.19.

Please cite as:

Williams PAH, Lovelock B, Cabarrus T, Harvey M

Improving Digital Hospital Transformation: Development of an Outcomes-Based Infrastructure Maturity Assessment Framework

JMIR Med Inform 2019;7(1):e12465

URL: <http://medinform.jmir.org/2019/1/e12465/>

doi: [10.2196/12465](https://doi.org/10.2196/12465)

PMID: [30632973](https://pubmed.ncbi.nlm.nih.gov/30632973/)

©Patricia AH Williams, Brendan Lovelock, Tony Cabarrus, Marlon Harvey. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 11.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Impact of an Electronic Patient Bedside Observation and Handover System on Clinical Practice: Mixed-Methods Evaluation

Alexandra Lang¹, PhD; Mark Simmonds², MRCP; James Pinchin³, PhD; Sarah Sharples³, PhD; Lorraine Dunn², BA; Susan Clarke², MSc; Owen Bennett², MSc; Sally Wood², MSc; Caron Swinscoe⁴, BSc

¹Trent Simulation and Clinical Skills Centre, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom

²Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom

³Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom

⁴NHS Digital, Leeds, United Kingdom

Corresponding Author:

Alexandra Lang, PhD
Trent Simulation and Clinical Skills Centre
Nottingham University Hospitals NHS Trust
Queen's Medical Centre
Derby Road
Nottingham, NG7 2UH
United Kingdom
Phone: 44 07921912376
Email: alexandra.lang@nottingham.ac.uk

Abstract

Background: Patient safety literature has long reported the need for early recognition of deteriorating patients. Early warning scores (EWSs) are commonly implemented as “track and trigger,” or rapid response systems for monitoring and early recognition of acute patient deterioration. This study presents a human factors evaluation of a hospital-wide transformation in practice, engendered by the deployment of an innovative electronic observations (eObs) and handover system. This technology enables real-time information processing at the patient’s bedside, improves visibility of patient data, and streamlines communication within clinical teams.

Objective: The aim of this study was to identify improvement and deterioration in workplace efficiency and quality of care resulting from the large-scale imposition of new technology.

Methods: A total of 85 hours of direct structured observations of clinical staff were carried out before and after deployment. We conducted 40 interviews with a range of clinicians. A longitudinal analysis of critical care audit and electronically recorded patient safety incident reports was conducted. The study was undertaken in a large secondary-care facility in the United Kingdom.

Results: Roll-out of eObs was associated with approximately 10% reduction in total unplanned admissions to critical care units from eObs-equipped wards. Over time, staff appropriated the technology as a tool for communication, workload management, and improving awareness of team capacity. A negative factor was perceived as lack of engagement with the system by senior clinicians. Doctors spent less time in the office (68.7% to 25.6%). More time was spent at the nurses’ station (6.6% to 41.7%). Patient contact time was more than doubled (2.9% to 7.3%).

Conclusions: Since deployment, clinicians have more time for patient care because of reduced time spent inputting and accessing data. The formation of a specialist clinical team to lead the roll-out was universally lauded as the reason for success. Staff valued the technology as a tool for managing workload and identified improved situational awareness as a key benefit. For future technology deployments, the staff requested more training preroll-out, in addition to engagement and support from senior clinicians.

(*JMIR Med Inform* 2019;7(1):e11678) doi:[10.2196/11678](https://doi.org/10.2196/11678)

KEYWORDS

health information technology; early warning score; mobile health; staff workload; clinical deterioration; patient safety; mixed methods

Introduction

Background

Patient safety literature has long reported the need for early recognition of deteriorating patients, with health strategy stating it as a health care priority [1-4]. Early warning scores (EWSs) are commonly implemented as “track and trigger,” or rapid response systems for monitoring and early recognition of acute patient deterioration.

We have presented an evaluation of a hospital-wide transformation from paper-based patient observations to mobile handheld device-enabled electronic observations (eObs) and electronic handover (eHandover) data collection in a large UK teaching hospital. Over 5500 mobile devices were deployed to over 6000 staff across 80 wards. The innovative technology allows real-time, automatic information processing at the patient’s bedside, with the aim of improving the efficacy of EWSs in practice and provide greater visibility of key patient data.

This study evaluated the deployment using objective measures and subjective evaluation of changes in clinical practice in addition to the overview of benefits realization based on analysis of hospital data sets.

Independent human factors researchers evaluated the technology implementation through pre- and postdeployment observations and staff interviews. Data extracted from the eObs and eHandover software and other supporting information technology (IT) systems provide evidence of the impact of this intervention on admissions to critical care and other benefits to the organization.

The combination of these datasets provides valuable insight into how a health information technology (HIT) intervention has affected care provision and patient safety in a large UK teaching hospital.

Background to Paper-Based Early Warning Scores and the Move to Electronic Observations

EWS systems deliver a standardized approach to observation frequency and response based around an aggregated scoring system which characterizes a patients’ physiological acuity. This process involves physiological observations being carried out at the patient’s bedside and a score being calculated (afferent limb). If this score meets the defined criteria, representing a significant abnormality, the observations are communicated and acted upon by appropriate clinical team members (efferent limb) [5]. For the system to be effective, both the afferent and efferent limbs need to be efficient, reliable, and timely. The system enables a proactive approach to assessment and recognition of the deteriorating patient, leading to reductions

in critical care admissions, mortality, and serious adverse events [6,7].

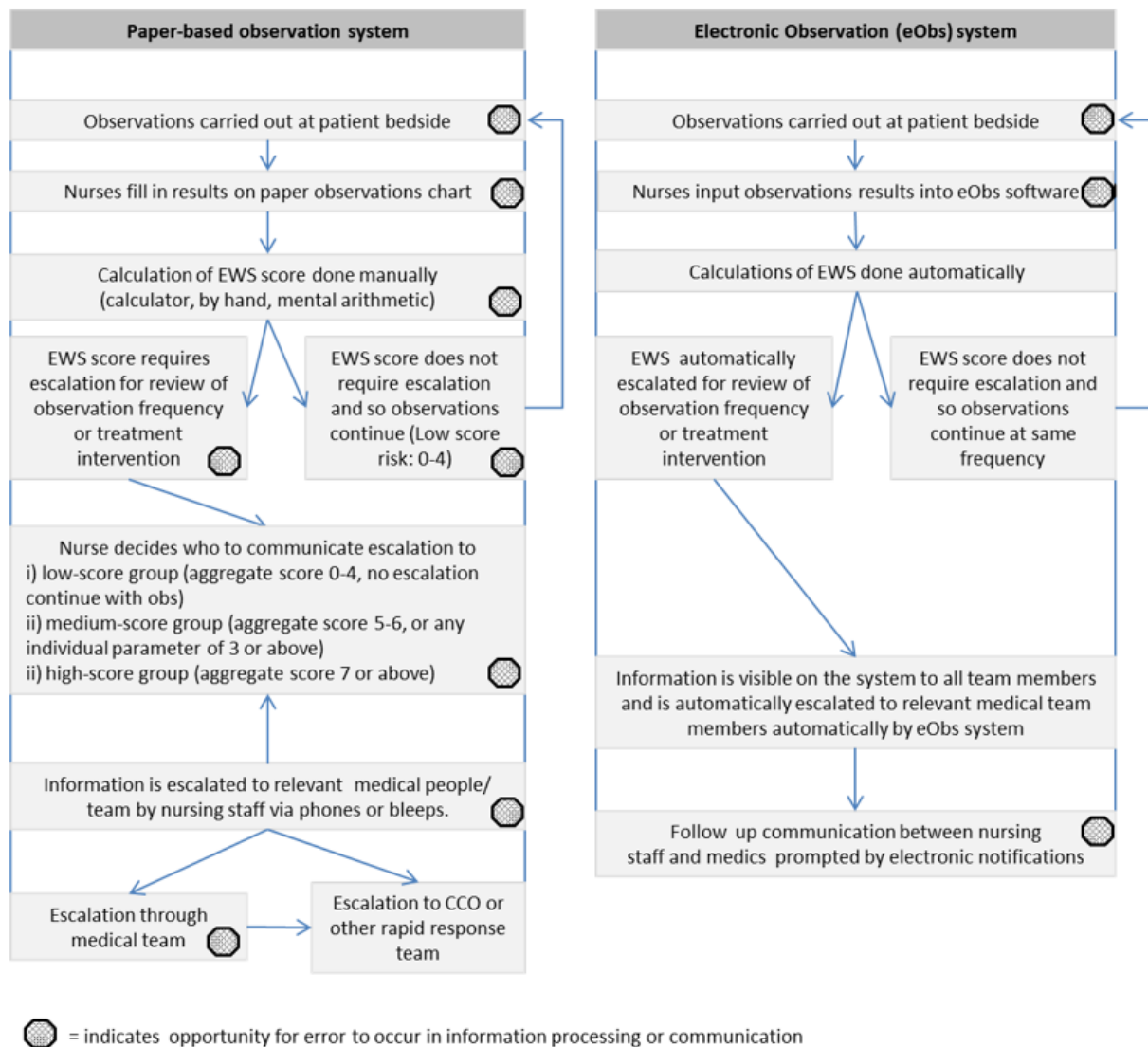
Historically, EWS processes have been implemented via paper-based charts; however, there is widespread acknowledgement of a variety of contextual reasons which lead to poor adherence to this practice [8,9]. To tackle poor adherence and associated organizational issues, Nottingham University Hospitals (NUH) National Health Service (NHS) Trust commissioned and facilitated the development of an electronic system (Nervcentre Software LtdTM), moving to mobile device access for all clinical staff.

Figure 1 shows the relative complexity of the paper-based system, demonstrating the risk of increased workload and reliance on active communication required of staff in a minimum of 8 decision points during EWS procedures. In comparison, the eObs system has fewer task stages and fewer interpersonal interactions, with automatic system actions replacing the manual score calculation tasks and decision points. To improve adherence, the eObs system also provides a reminder function for the next observation set. The higher number of tasks within the paper-based process increases the risk of potential errors at each stage of processing and communication of information as evidenced by Prytherchy et al [10]. This finding is well established in the literature on errors [11] and specifically the types of errors that occur in paper-based EWSs [12]. There is an increased risk of communication delay, as staff have to prioritize escalation of patient deterioration over other competing tasks. In contrast, the functions of the new electronic system streamline the process and reduce the number of opportunities for degradation of information.

The implementation of the eObs technology in 2015 was distinct from the EWS patient management policy which has been established on paper across NUH since 2008. By aligning precisely to the existing policy, a raised EWS is automatically and immediately escalated to senior clinical staff or the critical care outreach team (CCOT) through mobile instant messaging. The data recorded in the eObs module include all the physiological parameters previously calculated in the paper-based observation charts [13]. The system also allows for “special circumstances models” to be implemented where patient needs differ from standard EWS algorithms (eg, End of Life or known chronically abnormal physiology).

Equally, handover documents have historically been handwritten, nonstandardized, and at risk of being out-of-date, or incorrect, putting patient information and safety at risk [14]. The eHandover solution created a mobile platform to record key patient data in a standardized format, allowing different staff groups to access information where and when they need it. This functionality in eObs and eHandover provides opportunity for consistency, simplicity, and a reduction in the potential for perceptual error.

Figure 1. Visual comparison of clinical observation process using paper-based charts (left-hand flow diagram) and the electronic observations system (right-hand flow diagram). eObs: electronic observations; EWS: early warning score; CCO: critical care outreach.



Study Context

The study was carried out at a large teaching hospital trust in the United Kingdom, which provides secondary care services for approximately 2.5 million residents and facilitates approximately 1900 beds. Over a period of 9 months, personal mobile devices (iPhones and iPads) and training were provided to over 6000 nurses, doctors, health care assistants, and allied health professionals. The deployment was undertaken by a specialist Clinical Information and Communication Technology (CICT) team. This involved senior and practicing clinical nurses being developed into specialist Information and Communications Technology (ICT) advocates to lead the technology roll-out. Their dual role enabled them to support staff clinically while the general workforce were introduced to and becoming familiar with the eObs system.

This team coordinated the roll-out across 70 adult and children's wards at 2 hospital sites.

Methods

Study Design

The study collected pre- and postdeployment data concerning ward-based work via direct observation of staff before and after the deployment. Interviews and focus groups collected qualitative staff insights into the impact of mobile handheld devices and eObs on nursing and medical practice. Additional data sets were collected from existing hospital systems to give insight into the wider implications of eObs.

Ethical approval was obtained from an appropriate local ethics committee as a service evaluation project.

Structured Observations

Recruitment of clinical staff was done via flyers and facilitated introductions by the CICT team. Direct structured observations of clinicians were carried out pre- and postdeployment to record staff activities and location within the ward. A total of 23 predeployment and 64 postdeployment (n=87) observation sessions were obtained over 85 hours. Observation sessions

lasted between 15 and 120 min. Observers “shadowed” staff, using a bespoke tablet computer app to record activities and locations from exhaustive and exclusive lists. Researchers were not in attendance at the patient’s bedside but observed from a distance, and participants were informed that the observation could be halted at any time.

Observation sessions were divided into 30-second time bins. If an activity was observed in a 30-second bin, it was recorded as 1 observation even if multiple instances of the activity occurred (Figure 2). This method makes the observation of multitasking or rapid task switching possible and provides a measure of the relative distribution of different activities during the observation period.

Predeployment data were collected on a short stay acute medical admissions ward (n=11) and a health care of older person ward (n=13). Postdeployment data were collected on acute medical admissions wards (n=37), medical wards (n=12), and surgical wards (n=15).

Registered nurses were observed for a total of 17 hours predeployment (n=16) and 23.3 hours postdeployment (n=18). Doctors were observed for 10 hours predeployment (n=7) and 35.1 hours postdeployment (n=47). The participating doctors ranged in experience from consultant to Foundation Grade 1 (F1) doctor, which is the general postgraduate bridge between

medical school and training for full registration as a medical professional in the United Kingdom.

Interviews

Recruitment of staff for interviews was carried out via email and poster communications. Semistructured interviews and focus groups explored the impact of technology deployment on personal working practices and also encouraged reflection on the impact on teams, environment, and organization.

A total of 40 interview participants were recruited across a range of nursing and medical roles amounting to 18.5 hours of interview data. The number of interview participants for each staff type is given in Table 1. All interviews were carried out post system deployment, with the staff experience of eObs ranging from 1 week to 5 months.

Impact Evaluation

A longitudinal analysis of unplanned critical care admissions was derived from the NUH critical care audit dataset. Bed day costs were derived from local single organ high dependency unit (HDU) and 3 organ intensive care unit (ICU) support tariffs.

Review of EWS-related incidents on eObs wards was performed by 2 reviewers independently, from electronically recorded patient safety incident reports (Datix Ltd) from April 2014 to December 2015.

Figure 2. Encoding activity using the structured observation methodology. In (a), 8 sequential activities occur during the 1-min observation window. In (b), a set of unique activities is recorded for each of the two 30-second bins in the observation session. Note that “Looking at Notes” is only recorded once in the first bin despite 2 instances occurring in the first 30 seconds of observation. PC: personal computer.

Time	Category
0h00m7s	Talking on phone
0h00m12s	Looking at notes
0h00m24s	Using PC
0h00m26s	Looking at notes
0h00m30s	Using PC
0h00m35s	Looking at notes
0h00m45s	Walking within ward
0h00m52s	Talking face-to-face

(a)

➔

Bin number	Category
1	Talking on phone Looking at notes Searching for paperwork
2	Using PC Looking at notes Walking within ward Talking face-to-face

(b)

Table 1. Stratification of staff interviews.

Clinical role	Number of interviews	Total
Medical staff		
Consultants	5	18
Registrars	5	
Locums	1	
Junior doctors	4	
CCOT ^a	3	
Nursing staff		
Senior nursing staff	4	12
Staff nurses	6	
Health care assistants	2	
CICT ^b team	7	10
Critical skills educator	1	
Ward managers	1	
Hospital play specialists	1	
Grand total		40

^aCCOT: critical care outreach team.

^bCICT: clinical information and communications technology team.

Results

Structured Observations

Nurses

As expected, an increase in observations of “using smartphone” was detected after the deployment of the eObs technology (2.2% to 6.4% of 30-second bins). However, this change was small when compared with the reduction in time spent interacting with notes and talking on the phone. [Table 2](#) summarizes the changes for key observation categories [15].

A change was also seen in the observed location of nurses undertaking these activities. In particular, a move from the “office” (40.8% to 16.2% of the observed period) to the “nurse’s station” (13.3% to 35.1%) was observed.

A decrease in the number of activities observed in each 30-second bin was observed. This decrease is closely related to a decrease in rapid task switching. The mean number of

activities in each 30-second bin decreased from 1.99 (SD 0.04) to 1.66 (SD 0.03).

Doctors

The observed changes in the way doctors spend their time were similar to nurses. Smartphone use increased (3.7% to 8.3%) while remaining low relative to the frequency with which interacting with paper notes or desktop PCs was observed. Doctors were also observed spending less time in the office (68.7% to 25.6%) with more time at the nurses’ station (6.6% to 41.7%). Patient contact time more than doubled (2.9% to 7.3%).

One-Hour Case Example

[Table 3](#) presents an illustration of how the use of eObs and handheld mobile devices has changed working tasks and locations for clinicians. This example uses a 1-hour exemplar and assumes that each observed activity spanned the entire 30-second observation bin.

Table 2.

Observation category	Predeployment (% of 30-s observation bins)	Postdeployment (% of 30-s observation bins)
Using Personal Computer (PC) or Computer On Wheels (COW)	23.3	5.1
Looking at notes	36.2	22.3
Writing on notes	26.3	16.0
Talking on phone	8.4	4.0
Using smartphone	2.2	6.4

Table 3. One-hour example of changes in doctors and nurses' clinical tasks and locations of working in the ward owing to electronic observations.

Clinical role	Time
Doctors	
Task	
Average smartphone use	Increase from 2 min to 5 min
Location	
Time spent in office	Reduce from over 40 min down to less than 16 min
Time spent at the nursing station	Increase to 21 min
Time spent with patient	Increase from less than 2 min to over 4 min
Nurses	
Task	
Average smartphone use	Increase from just over 1 min to nearly 4 min
Use of personal computer (PC) or Computer On Wheels (COW)	Reduce from 14 min down to 3 min
Looking at notes	Reduce from over 21 min down to less than 14 min
Writing in notes	Reduce from nearly 16 min down to under 10 min
Talking on phone	Reduce more than half from 5 min to just over 2 min
Searching tasks	Reduce by 3.5 min
Location	
Time spent in office	Reduce from over 24 min to less than 10 min
Time spent at the nursing station	Increase from 8 to 21 min

Interviews

A total of 40 staff provided feedback about their experiences of the deployment process, eObs, and mobile devices.

The formation of a specialist clinical team who were trained as ambassadors to lead the deployment was universally lauded as a reason for the successful roll-out. This CICT team was praised for their capacity to multitask, assisting people with the technology while administering clinical care.

During the deployment and early use of the technology, staff reported increased stress and workload, with participants identifying a need for more training in advance of the deployment; however, this was fairly short-lived:

Not all of us had physically got our phone in time so it were all faffing, trying to get the phones charged up and all that kind of technical stuff...we'd not really had time to play with them...I think every one of us felt nervous the morning of it coming and I don't think we needed to. [Senior nurse]

The accessibility of information on the mobile devices appears to have streamlined staff discussions to quickly address treatment pathways by facilitating remote decision making and distributed working for both nursing and medical staff:

The best thing about it is it's a good record to consult, and certainly patients who've had multiple admissions, you can easily go back and see that information from previous admissions, and, unlike paper, it doesn't get lost. [Junior doctor]

Frustration was expressed by junior nursing and medical staff at a perceived lack of engagement with the new system by senior medical personnel (specifically consultants). It was considered that this issue was one of the main barriers to realizing the potential benefits in a ward setting. Several rationales were offered by medical staff (including consultants) to explain the lack of engagement by senior medical staff in the eObs deployment, including the perceived loss of expertise because of the "step change" in practice, the potential for embarrassment associated with use of the new system, or a general reluctance to embrace change.

Junior personnel (medics and nurses) provided an important source of informal device use support to individuals who were struggling to adopt the new system. This support was provided early on and during the weeks and months following the departure of the CICT team from the wards:

If you break it [eObs or phone], even now, it's a standard joke, we get one of the young staff to fix it. [Senior nurse]

"Word of mouth" or "heard it through the grapevine" communications often perpetuated information about eObs and device use throughout the workforce. The staff believed that if this "good practice" could be formally captured and disseminated, it could speed up the rate at which staff experienced benefits from the new system.

Nurse interview data revealed a largely positive response, reporting added value in the form of reassurance of patient health state owing to the real-time eObs information and also awareness of ward capacity. However, there was also an initial perception that the new technology could result in a loss of

control for nursing staff and promote a “Big Brother” culture, as automatic escalations meant taking away nursing autonomy. This perspective subsided over time as the real-time automation began to accelerate communications:

Mobile technology has made a huge difference to our working lives. It helps us to manage our workload and feel more in control of what is happening on the ward. It has reinforced the importance of communication between clinicians and has really demonstrated how patient care can be improved.

[Ward sister]

Over time, nursing staff began to identify how the system could potentially alleviate stress through greater visibility of information. Nurses began to describe the mobile devices as their own “personal tool” for workload management and improved awareness of team capacity:

It is just about making the device work for you...as I've got more confident with the device, I have said to my staff, just don't let it rule your shift and you get it to work for yourself.

[Deputy sister]

Nursing staff also described the value of handheld devices as a communication tool for use with patients and relatives, whereby the request for information could often be responded to more quickly.

For medics, initial access settings within the eObs system were at odds with current practice in terms of the perception about “consultant-led practice” versus the reality of registrars working independently. During the early deployment of eObs, relationships between consultants and registrars were put under pressure because of the permissions programmed within the system. Clinicians understood the need for policy to underpin the system; however, there was disruption to working practices as these issues were experienced:

...who delivers the cardiac arrest process and decides, well it's the registrar...so if you are allowing them to make those decisions then to say they can't alter the parameters is patronising. And it's again where the trust says 'all our decisions are consultant made' but the reality is that's not true.

[Consultant]

Medical staff explained how they used the device to “checkup” on patients that they had treated, for clinical reassurance, when they were off duty and had physically left the ward. It was acknowledged that this use of the device should not promote unhealthy practices with regard to work-life balance:

I think it is mainly when I see somebody sick in the ward when on acute medicine...and I just sneak a peek to make sure they are getting better instead of worse... [Registrar]

Medical staff believed that the new system had reduced time spent searching within a ward and had facilitated time management, a finding which is consistent with the observation data captured. This related not only to their working practices but also in their reflection of liaison with nursing staff and general hospital organization.

Longitudinal Impact Data

The roll-out of eObs has been associated with an approximate 10% reduction in total unplanned admissions to critical care units at NUH from eObs-equipped wards. No substantial change in hospital or critical care bedstock has occurred over this period. This benefit is more marked when critical care level and length of stay are taken into account (Figure 3). Alongside ongoing efforts to improved detection and response to the deteriorating patient [16], the impact of “real-time” communication of EWS and accessibility of patient information via handheld devices appears to be associated with a lower rate of critical care admissions, because of patients receiving more timely care and not requiring referral to those specialist services.

On the basis of nominal reference costs of £800 per Level 2 (HDU) bed day and £1200 per Level 3 (ICU) bed day, this 10% reduction in critical care admissions equates to an approximate £250K saving per quarter (Figure 4) since deployment.

More detailed cost breakdowns are required to understand the full economic cost-benefits of the scheme for the future, particularly with regard to the cost of the technology infrastructure (maintenance and replacement) against clinical health economic gains and prolonged use behavior of the system over time.

The eObs deployment is also associated with an approximate 50% reduction in reported EWS policy-related patient safety incidents in eObs wards. No such reduction was seen in non-eObs wards or in incident reporting in general over the same time period. Audit results also indicate adherence with EWS policy has improved because of the functionality of the system, namely, automated calculations, observation frequency setting, and user prompts that supports the findings from previous enquiries into electronic observations [17].

Figure 3. Unplanned admissions to level 2 (high dependency unit) and level 3 (intensive care unit).

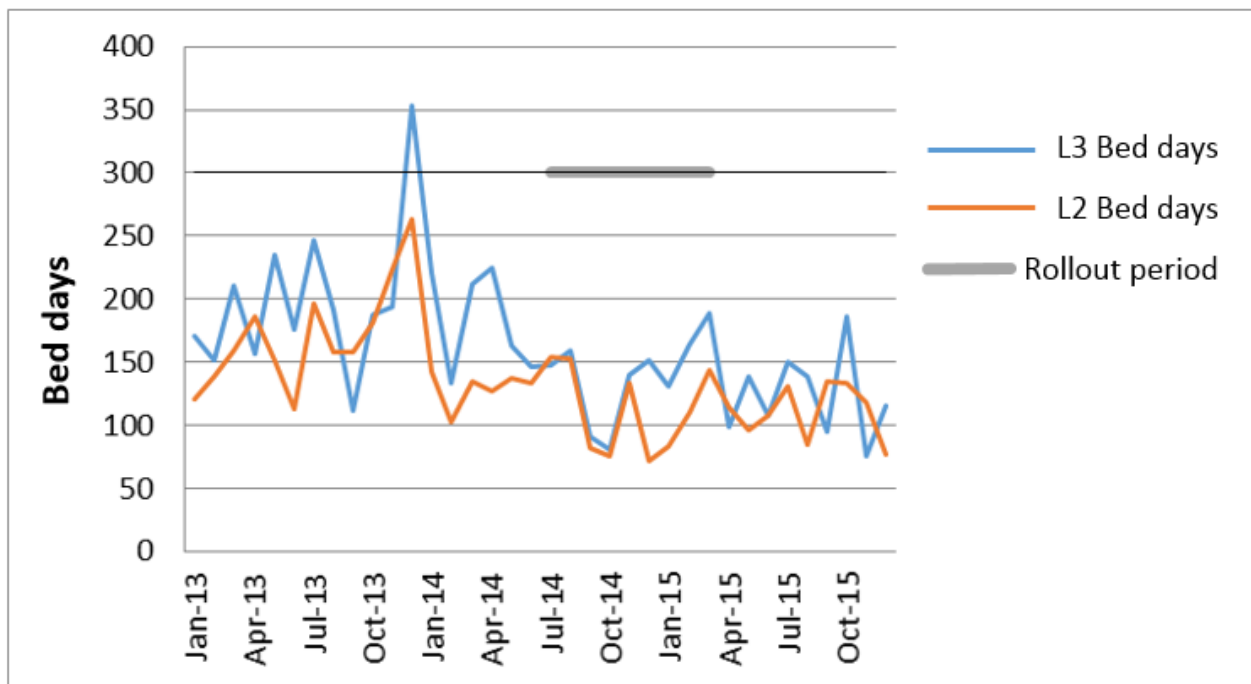
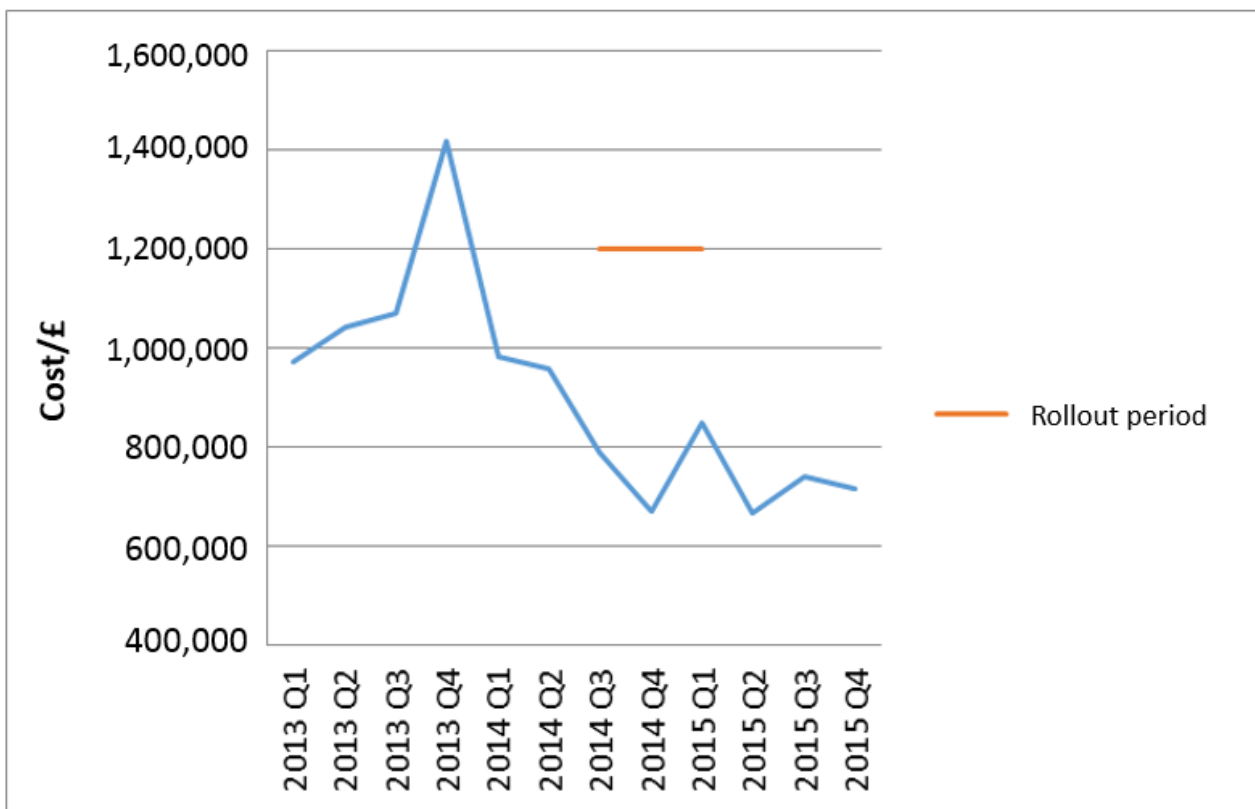


Figure 4. Cost of unplanned critical care admissions.



Discussion

Principal Findings

This evaluation has provided insight into the impact of mobile eObs and eHandover on working practice and elicited experiential data from the staff regarding their use of the new

systems. From these data, a range of benefits to the hospital trust and workforce has been identified. The interview and focus group data in particular have also indicated where additional research and development could further benefit staff and patient experience [15].

The mobile solution has reduced EWS-related patient safety incidents and has allowed nurses and doctors to spend more time with the patient at the bedside. Internal studies of the paper and eObs processes for taking and recording a full set of observations showed a time saving of 1 min 23 seconds per patient using eObs. On the basis of 7500 sets of observations taken at NUH each day, this equates to approximately 170 hours of nursing time saved every day, releasing time to care. This aligns with Stevenson's findings of how patient observations benefit from real-time capture at the point of care [18], and a reduction in nursing workload found by Wong et al [19].

Through personal ownership of devices, remote access has achieved real-time visibility of patient data across the whole hospital trust, allowing faster decision making and effective task prioritization. Clinicians are given vital information instantly because of automated escalations, and the need for multiple telephone calls is negated. This utility is further enhanced by users being able to access other medical apps and guidelines at the bedside. The eObs system appears to meet the strategies for EWS success identified by Russ *et al* [20] by being ubiquitous, being fit with ward workflow, and enabling records to be kept current and accessible.

This transformation of practice has made it easier for staff to “do the right thing” even when not in attendance at the bedside. The opportunities provided through remote, distributed working practices have achieved safer working (see Figures 3-5) while not compromising communication, as evidenced in the interviews. The time previously spent searching for paper and chasing colleagues—delays similar to those reported by Fox and Elliot [21] in their examination of a paper-based EWS system—has been replaced with more meaningful discussions based on the information now visible through the mobile interface.

Previous work has demonstrated other wider benefits such as indicating to hospital managers which wards are particularly busy [22]. Where eObs has provided improved transparency about team workload and ward capacity, staff and system can begin working together as a joint cognitive system, which in turn has supported the implementation of smart resource allocation in times of pressure.

The study demonstrates how different clinical roles interact in the uptake and success of changes in practice or technology interventions. The role of junior staff as informal mentors and early adopters of new practices and technologies was evident, while senior staff backing was seen as crucial to success but was perceived to be lacking in this instance.

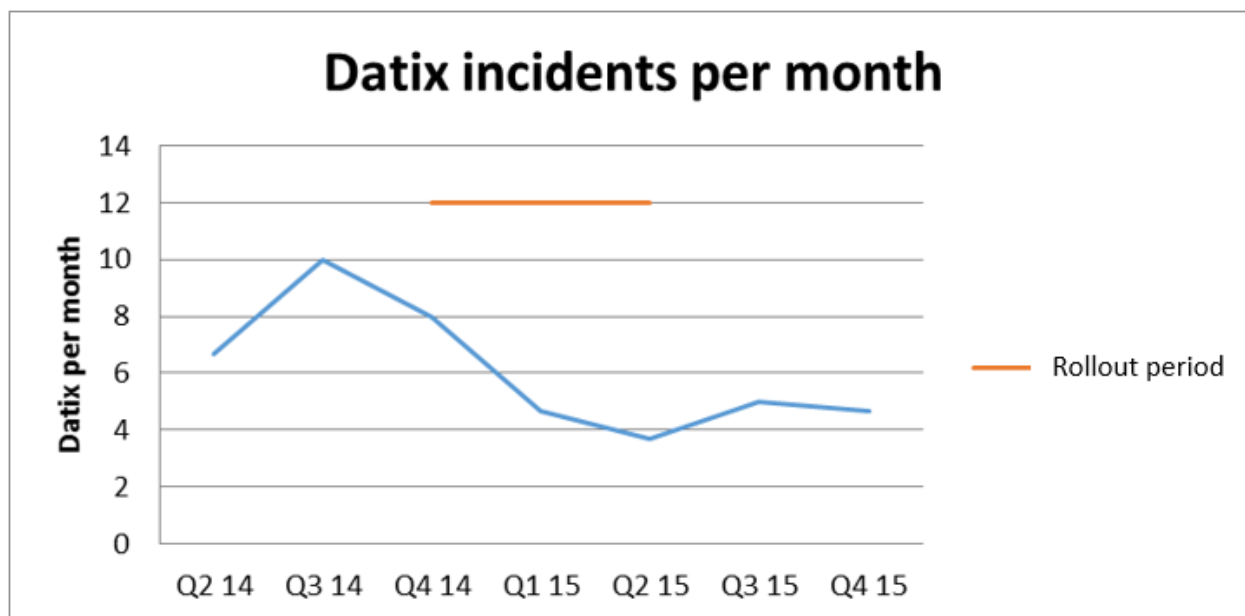
Challenges were also identified in technology integration and infrastructure, absence of feedback mechanisms for staff, management of expectations, and training requirements. Within

the scope of the technology deployment, infrastructure issues were continually being encountered, evaluated, and improved upon, for example, in regards to Wi-Fi “blackspots” which disrupted eObs operations. However, the staff felt that there was a lack of investment in ICT support during the critical roll-out period, specifically during out-of-hours shifts. The eObs training occurred very rapidly, and the staff felt that more time to understand the system functionality would have been valuable. There was undisputed praise for the facilitation of the CICT team in conducting the technology roll-out, providing vital technology assistance and staff ICT interface on the ward during deployment—a strategy that the hospital has learnt from and will likely implement again. The issues reported in the evaluation about the absence of a feedback loop were considered and, in response, the eObs operational team developed a more transparent and accessible way for staff to provide feedback to them and ICT support. These findings provide commentary for organizational learning regarding future technology deployments.

This study provided evidence to show that appropriately designed and deployed HIT can support improved situation awareness with regard to patient deterioration. By combining eObs (a frequently accessed utility) with an eHandover system, staff have become used to entering data on mobile devices and are contributing to team-held data on clinical, patient safety, and operational issues. This model of HIT use and deployment is one which could assist future technology deployments in other hospitals and in doing so support the work of Cresswell *et al* [23] and Greenhalgh *et al* [24] in improving HIT implementation.

This research establishes 8 principles of good practice which can contribute to successful HIT deployments and which have been realized through this study.

1. Mobile tools to support clinical observation have the potential to be beneficial for doctors and nurses.
2. Deployment of this technology takes time, must involve working with users, and must be supported by a specialist technology deployment team.
3. More junior staff adapt to the technology particularly well.
4. Clinicians find ways of using this technology in conjunction with other tools to manage their work.
5. Embedded algorithms must take account of different specializations.
6. The technology can support clinical and patient communications.
7. It is vital that there is integration of new IT systems with existing systems.
8. The technology is only as good as the infrastructure that supports it.

Figure 5. Early Warning Score–related incidents per month by quarter.

Limitations

The study data were limited to a single UK NHS Hospital Trust; however, the trust in question covers 2 sites and is the United Kingdom's fourth largest acute trust.

This study and the deployment of the eObs system coincided with a rolling deteriorating patient improvement program at the host hospital trust. The direct observation period was also limited to 2 months postdeployment, and the data do not reveal if the staff settled back into old routines. As such, this should be the topic of a further enquiry to establish cause and effect in regards to the technology in isolation of other quality improvement initiatives.

There is the potential for response bias within the interview data. The views of those willing to participate ($n=40$) may not be representative of the wider workforce (approximately $n=14,500$). To combat this natural effect of the qualitative approach, the interviewer involved was an independent researcher not affiliated with the hospital, and the staff were sampled from a range of job roles, with varying levels of experience of patient bedside observations and data during a range of shifts.

Future Work

There is opportunity to study the impact of further appropriation and expansion of eObs and specifically eHandover modules in clinical practice; analyze and measure the impact of improved situation awareness which is afforded by the technology and how to harness that information for effective workforce deployment and operational planning; investigate how mobile devices are being used on a personal level and where different clinical jobs and roles find utility in the technology so that this may be capitalized on and support further innovation [15].

Recent studies highlight the opportunities around continuous physiological monitoring of patients [25,26], utilizing technologies that are commonly used in HDUs. These solutions

show positive results with regard to responding to patient deterioration but are costly and require detailed cost-benefit analysis to understand the health economic benefit of monitoring on such a large scale in hospitals. In relation to this type of personalized monitoring [27], eObs provides a potential step change, whereby the data gathered within the system could be exploited to understand trends within population and medical groups.

While this study focused on just 1 hospital in the United Kingdom, there is significant growth in this area, with hospitals in the United Kingdom [28] procuring through 2 main service providers, System C (previously Vitalpac)—supplying 26 NHS Trusts [29]—and Nervecentre 35 NHS Trusts, and also supplying health care providers in Sweden and Australia [30]. Electronic health records (EHRs) have a greater degree of maturity, with statistics from the United States suggesting that nearly as many as 80.5% of hospitals are using EHRs to some degree in their care provision [31]. Hospitals in other less-developed jurisdictions such as China [32] and India are also investing in those systems, with lessons learnt from India echoing some of the experiences examined in this study [33]. As such, there is much to gain from health care providers and manufacturers of these technology platforms in sharing the lessons learnt from such large-scale deployments to ease the transition from paper-based to electronic working and improve key outcomes with regard to quality measurement, staff performance, and patient experience and safety.

Conclusions

The eObs and eHandover project has effected transformational changes in patient safety at a large acute hospital, bringing benefit to both staff and patients. In the hands of clinical staff, handheld devices and appropriate clinical software have the potential to reduce costs associated with inpatient management while simultaneously empowering staff in their daily activities, improving patient safety and releasing time to care.

Looking ahead, the full scale of the benefits experienced by this hospital trust is only just beginning to mature, with their full extent being realized.

Acknowledgments

This body of work is thanks to the collaborative efforts of NUH Trust CICT team and the eObs project team. The authors acknowledge the support of Health & Social Care Information Centre in supporting the contractual obligations and realization of benefits associated with the “NHS England Safer Hospitals, Safer Wards Technology Fund” award (Ref: SHSW-0087). The authors of this report are also supported by the Horizon Digital Economy Hub EP/G065802/1.

Conflicts of Interest

None declared.

References

1. Berwick DM, Calkins DR, McCannon JC, Hackbarth AD. The 100,000 lives campaign: setting a goal and a deadline for improving health care quality. *J Am Med Assoc* 2006 Jan 18;295(3):324-327. [doi: [10.1001/jama.295.3.324](https://doi.org/10.1001/jama.295.3.324)] [Medline: [16418469](https://pubmed.ncbi.nlm.nih.gov/16418469/)]
2. National Patient Safety Agency. London, UK: National Patient Safety Agency; 2007 Dec. Safety First: One Year On URL: <https://webarchive.nationalarchives.gov.uk/20090330105439/http://www.patientsafetyalliance.scot.nhs.uk/files/npsasafetyfirstresponsereportweb.pdf?fileid=283> [accessed 2019-01-18] [WebCite Cache ID 75WDFd44X]
3. Darzi A. Department of Health, NHS. London, UK; 2008. High quality care for all London: briefing on NHS next stage review final report URL: http://www.nhshistory.net/High_Quality_Care_for_All.pdf [accessed 2019-01-22] [WebCite Cache ID 75c81jyE9]
4. Royal College of Physicians. London, UK; 2013. Putting Patients First: realising Francis' vision. Royal College of Physicians' response to the final report of the Mid Staffordshire NHS Foundation Trust Public Inquiry URL: <https://www.rcplondon.ac.uk/file/1405/download?token=sB9Ro-gH> [accessed 2019-01-18] [WebCite Cache ID 75WELRINY]
5. McNeill G, Bryden D. Do either early warning systems or emergency response teams improve hospital patient survival? A systematic review. *Resuscitation* 2013 Dec;84(12):1652-1667. [doi: [10.1016/j.resuscitation.2013.08.006](https://doi.org/10.1016/j.resuscitation.2013.08.006)] [Medline: [23962485](https://pubmed.ncbi.nlm.nih.gov/23962485/)]
6. Jansen JO, Cuthbertson BH. Detecting critical illness outside the ICU: the role of track and trigger systems. *Curr Opin Crit Care* 2010 Jun;16(3):184-190. [doi: [10.1097/MCC.0b013e328338844e](https://doi.org/10.1097/MCC.0b013e328338844e)] [Medline: [20305556](https://pubmed.ncbi.nlm.nih.gov/20305556/)]
7. Mathukia C, Fan W, Vadyak K, Biege C, Krishnamurthy M. Modified Early Warning System improves patient safety and clinical outcomes in an academic community hospital. *J Community Hosp Intern Med Perspect* 2015;5(2):26716 [FREE Full text] [doi: [10.3402/jchimp.v5.26716](https://doi.org/10.3402/jchimp.v5.26716)] [Medline: [25846353](https://pubmed.ncbi.nlm.nih.gov/25846353/)]
8. Mackintosh N, Rainey H, Sandall J. Understanding how rapid response systems may improve safety for the acutely ill patient: learning from the frontline. *BMJ Qual Saf* 2012 Feb;21(2):135-144. [doi: [10.1136/bmjqs-2011-000147](https://doi.org/10.1136/bmjqs-2011-000147)] [Medline: [21972419](https://pubmed.ncbi.nlm.nih.gov/21972419/)]
9. Wood SD, Candeland JL, Dinning A, Dow S, Hunkin H, McHale S, et al. Our approach to changing the culture of caring for the acutely unwell patient at a large UK teaching hospital: A service improvement focus on Early Warning Scoring tools. *Intensive Crit Care Nurs* 2015 Apr;31(2):106-115. [doi: [10.1016/j.iccn.2014.11.006](https://doi.org/10.1016/j.iccn.2014.11.006)] [Medline: [25604030](https://pubmed.ncbi.nlm.nih.gov/25604030/)]
10. Prytherch DR, Smith GB, Schmidt P, Featherstone PI, Stewart K, Knight D, et al. Calculating early warning scores--a classroom comparison of pen and paper and hand-held computer methods. *Resuscitation* 2006 Aug;70(2):173-178. [doi: [10.1016/j.resuscitation.2005.12.002](https://doi.org/10.1016/j.resuscitation.2005.12.002)] [Medline: [16806641](https://pubmed.ncbi.nlm.nih.gov/16806641/)]
11. Reason J. Human error: models and management. *Br Med J* 2000 Mar 18;320(7237):768-770 [FREE Full text] [Medline: [10720363](https://pubmed.ncbi.nlm.nih.gov/10720363/)]
12. Clifton DA, Clifton L, Sandu DM, Smith G, Tarassenko L, Vollam SA, et al. 'Errors' and omissions in paper-based early warning scores: the association with changes in vital signs--a database analysis. *BMJ Open* 2015 Jul 3;5(7):e007376 [FREE Full text] [doi: [10.1136/bmjopen-2014-007376](https://doi.org/10.1136/bmjopen-2014-007376)] [Medline: [26141302](https://pubmed.ncbi.nlm.nih.gov/26141302/)]
13. Royal College of Physicians. London, UK: Royal College of Physicians; 2017 Dec 19. National Early Warning Score (NEWS) 2: Standardising the assessment of acute illness severity in the NHS - Report of a working party. Royal College of Physicians URL: <https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2> [accessed 2019-01-18] [WebCite Cache ID 75WFKUnQ4]
14. Raduma-Tomás MA, Flin R, Yule S, Williams D. Doctors' handovers in hospitals: a literature review. *BMJ Qual Saf* 2011 Feb;20(2):128-133. [doi: [10.1136/bmjqs.2009.034389](https://doi.org/10.1136/bmjqs.2009.034389)] [Medline: [21209133](https://pubmed.ncbi.nlm.nih.gov/21209133/)]
15. Lang AR, Pinchin J, Sharples S, Shaw D. Experiences of clinical staff during the introduction of mobile technology systems in acute care wards. In: Proceedings 19th Triennial Congress of the International Ergonomics Association (IEA). 2015 Presented at: 19th Triennial Congress of the International Ergonomics Association (IEA); August 9-14, 2015; Melbourne, Australia URL: <https://www.iea.cc/congress/2015/209.pdf>

16. Simmonds M. Royal College of Physicians. 2017 Feb 1. Your stories: improving patient safety - 'Recognise and Rescue programme': improving patients safety URL: <https://www.rcplondon.ac.uk/projects/outputs/your-stories-improving-patient-safety> [accessed 2019-01-18] [WebCite Cache ID 75WFjBbnZ]
17. Jones S, Mullally M, Ingleby S, Buist MI, Bailey M & Eddleston JM (2011). Bedside Electronic Capture of Clinical Observations and Automated Clinical Alerts to Improve Compliance with an Early Warning Score Protocol. *Critical Care and Resuscitation*, Vol. 13, No. 2 2011:83-88.
18. Stevenson JE. PhD thesis, University of Sheffield. 2016 Jan. Documentation of vital signs in electronic health records: a patient safety issue URL: <http://etheses.whiterose.ac.uk/12704/1/PhD%20Thesis%20Jean%20E%20Stevenson.pdf> [accessed 2019-01-22] [WebCite Cache ID 75cCEkBG6]
19. Wong D, Bonnici T, Knight J, Gerry S, Turton J, Watkinson P. A ward-based time study of paper and electronic documentation for recording vital sign observations. *J Am Med Inform Assoc* 2017 Jul 1;24(4):717-721. [doi: [10.1093/jamia/ocw186](https://doi.org/10.1093/jamia/ocw186)] [Medline: [28339626](https://pubmed.ncbi.nlm.nih.gov/28339626/)]
20. Russ AL, Saleem JJ, Justice CF, Woodward-Hagg H, Woodbridge PA, Doebbeling BN. Electronic health information in use: characteristics that support employee workflow and patient care. *Health Informatics J* 2010 Dec;16(4):287-305. [doi: [10.1177/1460458210365981](https://doi.org/10.1177/1460458210365981)] [Medline: [21216808](https://pubmed.ncbi.nlm.nih.gov/21216808/)]
21. Fox A, Elliott N. Early warning scores: a sign of deterioration in patients and systems. *Nurs Manag (Harrow)* 2015 Mar 25;22(1):26-31. [doi: [10.7748/nm.22.1.26.e1337](https://doi.org/10.7748/nm.22.1.26.e1337)] [Medline: [25806463](https://pubmed.ncbi.nlm.nih.gov/25806463/)]
22. DeVita MA, Smith GB, Adam SK, Adams-Pizarro I, Buist M, Bellomo R, et al. "Identifying the hospitalised patient in crisis"—a consensus conference on the afferent limb of rapid response systems. *Resuscitation* 2010 Apr;81(4):375-382. [doi: [10.1016/j.resuscitation.2009.12.008](https://doi.org/10.1016/j.resuscitation.2009.12.008)] [Medline: [20149516](https://pubmed.ncbi.nlm.nih.gov/20149516/)]
23. Cresswell K, Bates D, Sheikh A. Ten key considerations for the successful implementation and adoption of large-scale health information technology. *J Am Med Inform Assoc* 2013 Jun;20(e1):e9-e13 [FREE Full text] [doi: [10.1136/amiajnl-2013-001684](https://doi.org/10.1136/amiajnl-2013-001684)] [Medline: [23599226](https://pubmed.ncbi.nlm.nih.gov/23599226/)]
24. Greenhalgh T, Wherton J, Papoutsis C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017 Nov 1;19(11):e367 [FREE Full text] [doi: [10.2196/jmir.8775](https://doi.org/10.2196/jmir.8775)] [Medline: [29092808](https://pubmed.ncbi.nlm.nih.gov/29092808/)]
25. Zimlichman E, Szyper-Kravitz M, Shinar Z, Klap T, Levkovich S, Unterman A, et al. Early recognition of acutely deteriorating patients in non-intensive care units: assessment of an innovative monitoring technology. *J Hosp Med* 2012 Oct;7(8):628-633. [doi: [10.1002/jhm.1963](https://doi.org/10.1002/jhm.1963)] [Medline: [22865462](https://pubmed.ncbi.nlm.nih.gov/22865462/)]
26. Rothman MJ, Rothman SI, Beals J. Development and validation of a continuous measure of patient condition using the Electronic Medical Record. *J Biomed Inform* 2013 Oct;46(5):837-848 [FREE Full text] [doi: [10.1016/j.jbi.2013.06.011](https://doi.org/10.1016/j.jbi.2013.06.011)] [Medline: [23831554](https://pubmed.ncbi.nlm.nih.gov/23831554/)]
27. Capan M, Ivy JS, Rohleder T, Hickman J, Huddleston JM. Individualizing and optimizing the use of early warning scores in acute medical care for deteriorating hospitalized patients. *Resuscitation* 2015 Aug;93:107-112 [FREE Full text] [doi: [10.1016/j.resuscitation.2014.12.032](https://doi.org/10.1016/j.resuscitation.2014.12.032)] [Medline: [25597507](https://pubmed.ncbi.nlm.nih.gov/25597507/)]
28. Digital health. London UK: Digital Health Intelligence Limited Special Report: Observations vital signs URL: <https://www.digitalhealth.net/2018/07/special-report-observations-and-vital-signs-2/> [accessed 2019-01-18] [WebCite Cache ID 75WGpwJEO]
29. National Institute for Health and Care Excellence. 2016 Sep. VitalPAC for assessing vital signs of patients in Hospital. Medtech innovation briefing MIB79 URL: <https://www.nice.org.uk/advice/mib79> [accessed 2019-01-18] [WebCite Cache ID 75WHAC3KW]
30. Nervecentre. Wokingham UK: Nervecentre Software Ltd; 2018. Nervecentre - Our Customers URL: <http://nervecentresoftware.com/our-customers/> [accessed 2019-01-18] [WebCite Cache ID 75WHRzqAY]
31. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. *J Am Med Inform Assoc* 2017 Nov 1;24(6):1142-1148. [doi: [10.1093/jamia/ocx080](https://doi.org/10.1093/jamia/ocx080)] [Medline: [29016973](https://pubmed.ncbi.nlm.nih.gov/29016973/)]
32. Lei J, Sockolow P, Guan P, Meng Q, Zhang J. A comparison of electronic health records at two major Peking University Hospitals in China to United States meaningful use objectives. *BMC Med Inform Decis Mak* 2013 Aug 28;13(1):96 [FREE Full text] [doi: [10.1186/1472-6947-13-96](https://doi.org/10.1186/1472-6947-13-96)] [Medline: [23984797](https://pubmed.ncbi.nlm.nih.gov/23984797/)]
33. Scholl J, Syed-Abdul S, Ahmed LA. A case study of an EMR system at a large hospital in India: challenges and strategies for successful adoption. *J Biomed Inform* 2011 Dec;44(6):958-967 [FREE Full text] [doi: [10.1016/j.jbi.2011.07.008](https://doi.org/10.1016/j.jbi.2011.07.008)] [Medline: [21846508](https://pubmed.ncbi.nlm.nih.gov/21846508/)]

Abbreviations

CCOT: critical care outreach team

CICT team: Clinical Information and Communication Technology team

eObs: electronic observations

eHandover: electronic handover

EWS: early warning score
HDU: high dependency unit
HIT: health information technology
ICU: intensive care unit
ICT: intermediate care team
IT: information technology
NHS: National Health Service
NUH: Nottingham University Hospitals
PC: personal computer

Edited by G Eysenbach; submitted 31.07.18; peer-reviewed by A Bashir; comments to author 14.10.18; revised version received 12.11.18; accepted 22.11.18; published 06.03.19.

Please cite as:

Lang A, Simmonds M, Pinchin J, Sharples S, Dunn L, Clarke S, Bennett O, Wood S, Swinscoe C
The Impact of an Electronic Patient Bedside Observation and Handover System on Clinical Practice: Mixed-Methods Evaluation
JMIR Med Inform 2019;7(1):e11678
URL: <http://medinform.jmir.org/2019/1/e11678/>
doi: [10.2196/11678](https://doi.org/10.2196/11678)
PMID: [30839278](https://pubmed.ncbi.nlm.nih.gov/30839278/)

©Alexandra Lang, Mark Simmonds, James Pinchin, Sarah Sharples, Lorraine Dunn, Susan Clarke, Owen Bennett, Sally Wood, Caron Swinscoe. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 06.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Understanding Determinants of Health Care Professionals' Perspectives on Mobile Health Continuance and Performance

Ju-Ling Hsiao^{1*}, PhD; Rai-Fu Chen^{2*}, PhD

¹Department of Hospital and Health Care Administration, Chia-Nan University of Pharmacy and Science, Tainan City, Taiwan

²Department of Information Management, Chia-Nan University of Pharmacy and Science, Tainan City, Taiwan

* all authors contributed equally

Corresponding Author:

Rai-Fu Chen, PhD

Department of Information Management

Chia-Nan University of Pharmacy and Science

No 60, Sec 1, Erren Rd

Rende Dist

Tainan City, 71710

Taiwan

Phone: 886 6 2664911 ext 5308

Fax: 886 63660607

Email: rafuchen@gmail.com

Abstract

Background: With the widespread use of mobile technologies, mobile information systems have become crucial tools in health care operations. Although the appropriate use of mobile health (mHealth) may result in major advances in expanding health care coverage (increasing decision-making speeds, managing chronic conditions, and providing suitable health care in emergencies), previous studies have argued that current mHealth research does not adequately evaluate mHealth interventions, and it does not provide sufficient evidence regarding the effects on health.

Objective: The aim of this study was to facilitate the widespread use of mHealth systems; an accurate evaluation of the systems from the users' perspective is essential after the implementation and use of the system in daily health care practices. This study extends the expectation-confirmation model by using characteristics of individuals, technology, and tasks to identify critical factors affecting mHealth continuance and performance from the perspective of health care professionals (HCPs).

Methods: A questionnaire survey was used to collect data from HCPs who were experienced in using mHealth systems of a Taiwanese teaching hospital. In total, 282 questionnaires were distributed, and 201 complete and valid questionnaires were returned, thus indicating a valid response rate of 71.3% (201/282). The collected data were analyzed using WarpPLS version 5.0 (ScriptWarp Systems).

Results: The results revealed that mHealth continuance ($R^2=0.522$) was mainly affected by perceived usefulness, technology maturity, individual habits, task mobility, and user satisfaction, whereas individual performance ($R^2=0.492$) was affected by mHealth continuance. In addition, user satisfaction ($R^2=0.548$) was affected by confirmation and perceived usefulness of mHealth, whereas perceived usefulness ($R^2=0.521$) was affected by confirmation. This implied that confirmation played a key role in affecting perceived usefulness and user satisfaction. Furthermore, the results showed that mHealth continuance positively affected individual performance.

Conclusions: The identified critical factors influencing mHealth continuance and performance can be used as a useful assessment tool by hospitals that have implemented mHealth systems to facilitate the use and infusion of the systems. Furthermore, the results can help health care institutions that intend to introduce or develop mHealth applications to identify critical issues and effectively allocate limited resources to mHealth systems.

(*JMIR Med Inform* 2019;7(1):e12350) doi:[10.2196/12350](https://doi.org/10.2196/12350)

KEYWORDS

delivery of health care; mobile health; health information management; health care quality, access, and evaluation

Introduction

Background

With the widespread use of mobile technologies, mobile information systems (ISs) have become crucial tools in health care operations. In recent times, smart health (sHealth) has become a critical strategy that is promoted by the government and medical industry; however, the successful implementation of sHealth depends on the development of mobile health (mHealth) [1]. mHealth is defined as health care to anyone, anytime, and anywhere by removing location and temporal constraints while improving both the coverage and quality of health care [2].

Research Motivations and Purpose

In reality, health care professionals (HCPs) often require high-quality communication and information resources, including communication capabilities, hospital information systems (HISs), information resources, and clinical software applications, at the point of care to facilitate rapid decision making with a low error rate, improve the quality of data management and accessibility, and improve practice efficiency and knowledge [3-8]. Although the appropriate use of mHealth may result in major advances in expanding health care coverage, increasing decision-making speeds, managing chronic conditions, and providing suitable health care in emergencies [9], Solanas et al [1] argued that mHealth is still in its early stages of development. Mechael et al [10] reported that current mHealth research does not adequately evaluate mHealth interventions, and it does not provide sufficient evidence regarding the effects on health. In addition, the World Health Organization [11] indicated that competing priorities, cost, and lack of knowledge are the most crucial barriers to mHealth implementation. Thus, appropriate evaluations, specifically after the implementation of mHealth systems and their use in daily health care practice, are critical, particularly from users' (HCPs') perspectives. This paper proposes and validates an extended model by integrating the expectation-confirmation model (ECM) and the characteristics of individuals, technology, and tasks to identify critical factors affecting mHealth continuance and performance from the perspective of HCPs and assessing the infusion of mHealth in clinical practice.

Literature Review

Mobile Health

Varshney [2] defined mHealth as health care to anyone, anytime, and anywhere by removing locational and temporal constraints while improving both the coverage and quality of health care. Alternatively, mHealth is the application of mobile communication technology in the field of health care; it integrates HISs and mobile devices with wireless communication technologies to achieve immediate medical care and handle diverse cooperative medical tasks [12]. Nowadays, various mobile devices—personal digital assistants, tablet personal computers (PCs), notebook computers, personal handy-phone system, smartphones, panel PCs, mobile clinical assistants, and iPads—have been used in accessing mobile ISs through wireless networks in clinical settings owing to their portable size,

relatively low costs, and ease of use [13,14]. The term *mobile* emphasizes various abilities and conditions as well as movability and portability. Increasing mobility can also enhance service efficiency and flexibility. Previous studies have indicated that the need for mobility is the primary reason for the applications of technological innovations in hospitals, and mobility is crucial in health care [14-16]. Thus, mHealth has the potential to increase the speed, work quality, and efficiency of HCPs. The implementation of mHealth is often achieved using portable information devices, such as a tablet PC, notebook, iPad, or smartphone, to appropriately address the needs of HCPs.

Many studies have reported that when appropriately used, mHealth systems facilitate rapid decision making with low error rates, thereby improving the quality of data management and accessibility and improving practice efficiency and knowledge [3-8]. Some researchers [17-19] mentioned that mHealth systems improve the quality of health care services, increase the productivity of HCPs, and ensure the timeliness of information provision, thus reducing the occurrence of errors. Therefore, mHealth systems are expected to exert considerable effects on clinical routines and workflows.

Information Technology Continuance and Performance

Bhattacharjee [20] argued that existing information technology (IT) or information system (IS) acceptance models, focusing on user evaluations at the early stage of IT or IS adoption and implementation, provide an inadequate explanation of and may sometimes contradict observed continuance behaviors; moreover, the long-term success of an IT or IS depends on its continued use rather than its first-time use. Bhattacharjee [20] proposed an ECM, one of the earliest IS continuance models, based on expectation-confirmation theory [21] in consumer behavior for understanding IS continuance after implementation, where the use of ISs transcends conscious behavior and becomes part of the normal routine activity. The study revealed that users' willingness to continue using ISs was affected by user satisfaction and perceived usefulness after using ISs. Moreover, the expectation-confirmation and perceived usefulness of ISs directly affect IS users' satisfaction; user satisfaction directly affects the willingness to continue using ISs. Limayem et al [22] further suggested that information communication technology (ICT) implementation should be considered a success when a significant number of users progress from the initial adoption stage to using ICT on a continuing basis. Nowadays, the ECM is being widely used and extended to investigate factors affecting user intentions regarding ISs after IS implementation and behaviors in various research contexts, including Web portals [23], online communities [24], electronic medical records [25], mHealth systems [26], and e-service [27-30]. Among the aforementioned studies, Akter et al [26] considered that continuance is a challenge for mHealth systems and that exploring theories on continuance behavior is necessary for developing a comprehensive continuance model for understanding mHealth services. Thus, Akter et al [26] incorporated the ECM and the constructs of service quality and trust to investigate the continuance of mHealth services at the bottom of the economic pyramid. Mettler [25] integrated the ECM and factors affecting automatic behavior (facilitating conditions, task fit, and computer literacy) to evaluate electronic

medical record continuance behavior. Furthermore, Chen et al [29] investigated the effects of technology readiness (innovativeness, optimism, discomfort, and insecurity) on user satisfaction and continued intention of e-services. Although many extended ECM studies were conducted, Bhattacharjee and Barfar [31] argued that some studies are inappropriate to just integrate acceptance and continuance theory to predict IS continuance behavior. This implied that the extended ECM should consider some salient variables in the IS Infusion (assimilation or integration) stage that a specific IS has been well implemented and become a part of their daily routine processes.

Some studies have emphasized examining the determinants of mHealth in the assimilation or integration stage, where the mHealth services or systems are stable and have been incorporated into routine practices [17,32-34]. For example, O'Connor et al [33] argued that most infusion studies have paid considerable attention to the technological aspects at an organization level rather than at an individual level. The authors suggested that additional studies be conducted at the technology infusion stage at an individual level by considering characteristics of technology, individuals, and tasks. Therefore, they proposed a research framework that focused on investigating the effects of the characteristics of technology (availability, maturity, and portability), individuals (habits, self-efficacy, and technology trust), and tasks (time criticality, interdependence, and mobility) on the extent of the infusion of mHealth services by HCPs and the relationship between the extent of infusion (including integrative use and exploratory use) and performance based on the results of an in-depth case study. Although the study proposed many potential factors influencing mHealth systems in the infusion stage, the framework should be appropriately modified and validated according to various health care contexts or applications.

The performance of mHealth ISs should be evaluated based on user satisfaction and the specific outcomes of their continued use from users' perspectives as performance evaluation is a major concern of the effects of ITs or ISs [35-37]. Goodhue and Thompson [37] proposed the task-technology fit (TTF) theory to highlight the importance of the fit between the characteristics of technologies and user tasks in achieving the effects of individual performance. In the model, the TTF is affected by antecedents (including technologies and tasks characteristics), and the TTF also has a significant effect on IS utilization and performance. In addition, IS utilization also has a direct effect on individual performance in TTF theory. On the basis of the TTF perspective, Hsiao and Chen [32] found that the use of mobile ISs provided nursing staff with real-time and accurate information and increased their efficiency and effectiveness in patient-care duties, thus further improving nursing performance. Lin [17] reported a significant effect of the fit among technology (applicability, user interface, and portability), individual (computer self-efficacy, user experience, and self-immersion), and task (nonroutine, timeliness, interdependence, and mobility) characteristics on task performance (in terms of meeting

expectations, positive attitude, and meeting user needs) of mHealth systems. Although the TTF theory provides a useful perspective for investigating the relationships among TTF, utilization, and performance, there is a lack of empirical studies, particularly in the health care industry in exploring the role of TTF in the IS infusion stage about IS being integrative and exploratory used [38]. In this study, we only refer to the results of antecedents of TTF and the relationship between IS utilization and individual performance of health information technologies' applications.

Methods

Research Model

To provide comprehensive understanding and insights into the postimplementation stage of mHealth systems (or at IS infusion stage), this study proposed an extended ECM research model for investigating key factors affecting the continuance and performance of mHealth services in Taiwan by incorporating the ECM proposed by Bhattacharjee [20] and the framework of mHealth infusion proposed by O'Connor et al [33]. This integration is based on the assumption that mHealth continuance is critical in the IS infusion stage because the use of mHealth services has become a part of daily clinical practice for HCPs. Moreover, the continuance intentions of HCPs and the subsequent use behavior of mHealth systems are expected to enhance individual performance. However, some variables mentioned in the framework proposed by O'Connor et al [33] should be adjusted according to the health care contexts and applications in Taiwan. In this study, the variables of self-efficacy and technology trust of user characteristics in O'Connor et al [33] were excluded. Self-efficacy is not considered to be a significant factor in mHealth infusion because it is insignificant in studies of physicians' [39] and nurses' [40] HIS acceptance in Taiwan. Furthermore, the technology trust proposed by O'Connor et al [33] to address the problem that users may be reluctant to use the IT because of its reliability. In this study, the mHealth applications have been used in the case hospital for several years, and they have been infused into HCPs' daily clinical practices. In addition, mHealth applications are not mandatorily used by HCPs; therefore, the technology trust is not a major concern in this study. We also append personal innovativeness as an investigated factor in individual characteristics as Rai et al reported that consumers' personal innovativeness exerted significantly positive effects on mHealth usage intention and assimilation [34]. However, the relationship between personal innovativeness and mHealth continuance should be further validated in the health care contexts of Taiwan.

Therefore, the research model (Figure 1) can be divided into 2 major parts. The first part includes the aspects affecting electronic health (eHealth) continuance and performance derived from the ECM and the effects of mHealth continuance: confirmation, perceived usefulness, user satisfaction, continuance intention, and individual performance.

Figure 1. Research framework. H1: The confirmation of mHealth systems significantly affects perceived usefulness; H2: The confirmation of mHealth systems significantly affects user satisfaction; H3: The perceived usefulness of mHealth systems significantly affects user satisfaction; H4: User satisfaction with mHealth systems significantly affects mHealth continuance; H5: The perceived usefulness of mHealth systems significantly affects mHealth continuance; H6: The continuance of mHealth significantly affects individual performance; H7: The individual characteristics of HCPs significantly affect mHealth continuance; H8: The technology characteristics of mHealth significantly affect mHealth continuance; H9: The task characteristics of HCPs significantly affect mHealth continuance; mHealth: mobile health.

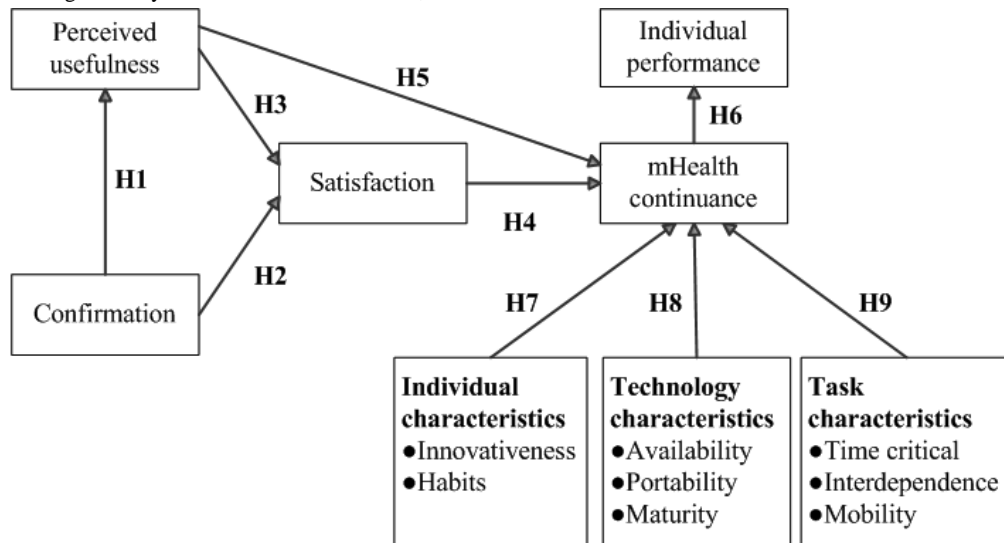


Table 1. Measurement and operational definitions of variables.

Construct	Operational definition	Source	Measurement items
Confirmation	Users' perception of the congruence between expectation of mHealth ^a use and its actual performance	[20]	4
Perceived usefulness	Users' perception of the expected benefits of mHealth use	[21,41]	5
User satisfaction	Users' affect with (feelings about) mHealth use	[20,42]	3
mHealth continuance	Users' intention to continue using mHealth	[20,25]	3
Individual			
Habits	The extent to which an individual tends to use the mHealth automatically	[22,33]	4
Innovativeness	Willingness to try out any new technology	[29,41]	4
Technology			
Availability	The ability of accessing patient information when required	[17,33]	3
Portability	The degree of ease associated with transporting the mHealth	[17,33]	3
Maturity	The existence of a level of system quality that is perceived as satisfactory and the perceived need for system improvement by the user.	[33,43]	3
Task			
Time critical	The urgency when accessing information through the mHealth	[17,33]	3
Interdependence	The degree to which completing tasks using mHealth requires interaction with other people	[17,32,33]	3
Mobility	The extent to which a task is being performed in different locations using the mHealth	[17,33]	3
Individual performance	The use of mHealth can help health care practitioner improve efficiency, effectiveness, and quality of medical activities	[33,44]	6

^amHealth: mobile health.

The second part investigates the effect of characteristics of technology, individual, and task on mHealth continuance; this part is based on the framework proposed by O'Connor et al [33], which involves innovativeness and habits (individuals), availability, portability, maturity (technology), time criticality,

interdependence, and mobility (tasks). The measurement, operational definition, and the number of items for the variables are summarized in Table 1.

In this study, confirmation refers to the users' perception of the congruence between expectation associated with the use of

mHealth systems and their actual performance. Perceived usefulness refers to the users' perception of the expected benefits of mHealth use. User satisfaction refers to users' affect (feelings) regarding previous mHealth use. As shown in the ECM, confirmation has a direct effect on perceived usefulness and user satisfaction [20]. Previous studies have argued that perceived usefulness directly affects user satisfaction and continuance of the use of system [20,25], and user satisfaction exerts direct effects on the continuance of the use of ISs [20,25,42]. Furthermore, mHealth continuance refers to the users' intention to continue using mHealth. Individual performance refers to the use of mHealth to help HCPs improve efficiency, effectiveness, and quality of medical activities. Individual performance implies an increase in efficiency and improvement in work efficacy and quality [37,44]. O'Connor et al [33] argued that users' continuance intentions could be affected by user satisfaction, and it can be used to predict or explain the continued use of IS as well as individual performance. Thus, 6 research hypotheses are proposed as follows:

- H1: The confirmation of mHealth systems significantly affects perceived usefulness.
- H2: The confirmation of mHealth systems significantly affects user satisfaction.
- H3: The perceived usefulness of mHealth systems significantly affects user satisfaction.
- H4: User satisfaction with mHealth systems significantly affects mHealth continuance.
- H5: The perceived usefulness of mHealth systems significantly affects mHealth continuance.
- H6: The continuance of mHealth significantly affects individual performance.

Most factors mentioned in the framework proposed by O'Connor et al [33] are suitable for mHealth continuance as continuance is critical in the mHealth infusion stage in which the use of mHealth services has become a part of daily clinical practice routine of HCPs. The influencing factors investigated in this study can be divided into 3 dimensions: individual characteristics (innovativeness and habits), technology characteristics (availability, maturity, and portability), and task characteristics (time criticality, interdependence, and mobility). The aforementioned individual characteristics represent individual traits and perceptions after using IT. Personal innovativeness is defined as a personal willingness to try using new information technologies [45], and it is a personal trait in technology-usage behavior [15,46]. Users with high innovativeness are often concerned with the development of new technologies, effectively learning technology functions on their own and providing suggestions to others. Rai et al [34] reported that consumers' personal innovativeness exerted significantly positive effects on mHealth usage intention and assimilation. Thus, high personal innovativeness facilitates changes in organizations and the diffusion of implemented technologies in internal operations [45]. Chen et al [29] reported that innovativeness is a major driver of user satisfaction, and users with higher innovativeness exhibit a relatively high level of satisfaction with technologies and the willingness to continue using them. Therefore, personal innovativeness significantly

affects mHealth continuance. Habit has been defined as the extent to which an individual tends to use the technology automatically [22]. A habit is often considered a reflective behavior or action taken without much consideration. With time, ideas, methods, judgments, and reactions stabilize. Such a behavioral model becomes partially fixed; thus, it is referred to as inertial. Habits are also reported to affect previous and future behavior [47]. Gefen [48] reported that the use of technology becomes a personal habit for users if the use of innovative technology is routine behavior. Habits strengthen a personal behavior through repeated stimulation and reaction. Limayem et al [22] noted that users who have used mHealth technologies for more than one year gain habits that involve using technologies. Thus, habits were expected to influence the infusion of mHealth technologies and are considered to significantly influence mHealth continuance.

Goodhue and Thompson [37] indicated that a better fit between task and technology characteristics improves performance in the infusion stage when technology has been used continuously. Therefore, task and technology characteristics should be considered while investigating technology performance. Task characteristics imply the inherent nature of tasks that users are expected to execute. In this study, task characteristics comprised time criticality, interdependence, and mobility. Time criticality is defined as urgency when accessing information through mHealth technology [33,49]. Interdependence is the degree to which completing tasks using mHealth technology requires interaction with other people [33,50]. Mobility is the extent to which a task can be performed in different locations using mHealth technology [33,49]. Technology characteristics are the specific features, functionality, or usability provided by specific technologies. Technology characteristics comprise availability, maturity, and portability. Availability is defined as the ability to access mHealth technologies when required [33,51]. Maturity is related to the existence of a level of system quality that is perceived as satisfactory and the perceived need for system improvement by users [33,43,52]. Portability is the degree of ease associated with transporting mHealth technologies [33,53]. As shown, 3 research hypotheses, including 8 subhypotheses, are proposed as follows:

- H7: The individual characteristics of HCPs significantly affect mHealth continuance.
 - H7a: Personal innovativeness significantly affects mHealth continuance.
 - H7b: Individual habits significantly affect mHealth continuance.
- H8: The technology characteristics of mHealth significantly affect mHealth continuance.
 - H8a: The availability of mHealth significantly affects mHealth continuance.
 - H8b: The portability of mHealth significantly affects mHealth continuance.
 - H8c: The maturity of mHealth significantly affects mHealth continuance.
- H9: The task characteristics of HCPs significantly affect mHealth continuance.

- H9a: Task time criticality significantly affects mHealth continuance.
- H9b: Task interdependence significantly affects mHealth continuance.
- H9c: Task mobility significantly affects mHealth continuance.

Instrument and Respondents

The questionnaire was designed in 2 stages. The first stage involved the establishment of measurement items. We collected results from literature reviews to obtain a comprehensive list of measurement items. All measures for each construct were obtained from existing validated instruments, and they were modified to ensure the appropriateness for mHealth. A total of 4 variables, namely confirmation, perceived usefulness, satisfaction, and mHealth continuance, which were derived from the ECM, were measured using 15 items adapted from Bhattacharjee [20], Mettler [25], Wright and Marvel [42], and Kuo et al [41]. The performance was measured using 6 items adapted from O'Connor et al [33] and Junglas et al [44]. Individual characteristics, including innovativeness and habits, were measured using 8 items adapted from Limayem et al [22], Chen et al [29], O'Connor et al [33], and Kuo et al [41]. Technology characteristics, including availability, portability, and maturity, were measured using 9 items adapted from Lin [17], O'Connor et al [33], and Gebauer et al [43]. Task characteristics, including time criticality, interdependence, and mobility, were measured using 9 items adapted from Hsiao and Chen [32], Lin [17], and O'Connor et al [33]. The detailed descriptions related to the survey questionnaire are provided in the [Multimedia Appendix 1](#). The questionnaire items were preliminarily translated into Chinese, and 2 experts in bilingual education in health care and information management were invited to evaluate the content equivalence of the translations. The questionnaire comprised 2 major parts. The first part collected participants' demographic data, including age, sex, education level, department, and experience in using mobile technologies and mHealth. The second part included measure items related to the factors influencing mHealth continuance and performance.

The second stage of questionnaire design involved the evaluation and selection of the measurement scale. A content validity index (CVI) was used to evaluate the questionnaire content according to a threshold value of 0.8 for item selection suggested by Petrick [54]. A total of 2 HCPs of mHealth and a professor of health informatics management were invited as experts to examine the content validity of the questionnaire. Among the initial questionnaire containing 47 items, except for 1 item, which was excluded as its CVI was less than .8, 46 items were retained as the CVI values were greater than .95 and the average CVI was .98, which indicated excellent expert validity. Furthermore, the semantics and wording of the questionnaire were revised according to experts' suggestions. Finally, a 46-item questionnaire was obtained. Each item was measured using a 5-point Likert scale (1 for strongly disagree and 5 for strongly agree).

The respondents of this study were the HCPs of the target hospital with approximately 120 doctors and 500 nurses in

southern Taiwan. Since 2009, the case hospital has developed and implemented mHealth systems, a combination of mobile ISs and medical devices, for satisfying HCPs' needs of clinical patient care, particularly for providing more timely communication of HCPs and direct data input at source, reducing possible medical errors, and accessing up-to-date medical records. The mobile ISs can connect and access all required and integrated patient-related information from hospital ISs, including various developed systems (computerized physician order entry system, laboratory ISs, nursing ISs, pharmacy ISs, picture archiving and communication system, electronic medical records, patient referral system, and others) for supporting inpatient, outpatient, and emergency services in a hospital, through a secure wireless network infrastructure. The mobile ISs can be installed on various mobile devices, including a mobile nursing cart equipped with a Tablet PC (specifically for nurses), mobile medical cart equipped with a Tablet PC (specifically for physicians), mobile phones, and iPad for satisfying the mobile needs of HCPs, particularly in the inpatient and emergency services.

Since 2014, some health apps of the case hospital have been developed and installed on mobile phones and iPads for providing instant access to the results of medical examinations, tests, and reports and receiving immediate notifications from high-risk reminder systems for clinical laboratory critical value alerts; however, those apps only provide relatively specific and limited information for patient care because of the limitations of small screen size, less computation power, and data key-in problems of intelligent mobile devices. Therefore, HCPs in the case hospital prefer accessing full patient care information through mobile ISs installed on the mobile nursing cart, mobile medical cart, and tablet PC. HCPs who had at least one year of experience in mHealth apps and were active and voluntary users of mHealth, using mobile ISs through mobile devices in clinical practices, were requested to participate. After obtaining approval from the Institutional Review Board (IRB NO.105B-009), the questionnaires were distributed to qualified HCPs under the assistance of the nursing department and hospital administration department. The duration of data collection was from February 1 to March 1 in 2016.

Results

Descriptive Statistics

The survey was administered to 282 respondents, and 201 valid responses were returned, which indicated a response rate of 71.3% (201/282). Voluntary participation might explain the relatively high response rate. The demographic data (see [Table 2](#)) showed that most of the respondents (94.0% [189/201]) were female, 92.1% (185/201) were less than 40 years old (48.3% [97/201] and 43.8% [88/201] were aged <30 years and 30-40 years, respectively), and 73.1% (147/201) had a bachelor's or master's degree. Among the respondents, 94.0% (189/201) worked in the nursing department, whereas the others worked in the medical department. Moreover, 77.1% (155/201) of the participants had more than 1 year of experience in using mHealth, thus indicating the appropriateness of the selected respondents.

Table 2. Demographic data (n=201).

Measure or category	Statistics
Age (years), n (%)	
<30	97 (48.3)
31-40	88 (43.8)
41-50	12 (6.0)
51-60	4 (2.0)
Gender, n (%)	
Male	12 (6.0)
Female	189 (94.0)
Education level, n (%)	
Junior college	54 (26.9)
Bachelor	144 (71.6)
Master (or higher)	3 (1.5)
Department, n (%)	
Medical (Physicians)	12 (6.0)
Nursing (Clinical nurses)	189 (94.0)
Experience in using mobile technologies (years), n (%)	
1-3	146 (72.6)
3-6	43 (21.4)
6-9	7 (3.5)
>9	5 (2.5)
Experience in using mobile health (years), n (%)	
1	46 (22.9)
1-5	145 (72.1)
5-10	10 (5.0)

Measurement Model

The collected data were analyzed using the partial least square (PLS) technique, which can offer extensive, scalable, and flexible causal-modeling capabilities [55], in the WarpPLS software (Version 5.0) because of its ease of use as well as its capability of performing all the modeling procedures reported in this study [56]. A 2-step approach of the PLS technique suggested by Chin [57] was used. The first step was to evaluate the measurement model, whereas the second step focused on evaluating the structural model. Several criteria are recommended for assessing the model-data fit when using WarpPLS 5.0, including average path coefficient (APC), the average R-squared (ARS), average adjusted R-squared (AARS), average block variance inflation factor (AVIF), average full collinearity variance inflation factor (AFVIF), Tenenhaus Goodness of Fit (GoF), and R-squared contribution ratio (RSCR) [56]. These model fit and quality indices are other advantages provided by WarpPLS 5.0 than other variance based structural equation modeling methods. In general, the addition of latent variables into a model will increase the value of ARS but decrease the value of APC. Both ARS and APC will increase

simultaneously only when the addition of latent variables can improve the overall model predicative and explanatory quality [56]. The AARS, generally lower than ARS in a model, is used to correct improper increases in R-squared coefficients when predictors cannot improve the explanatory value in each latent variable [56]. The AVIF and AFVIF are used to evaluate the increase of collinearity of the model if new latent variables are added and that may overlap in meaning with existing latent variables [56]. The GoF is a measure for evaluating the model's explanatory power, while RSCR is a measure evaluating the extent that a model is free from negative R-squared effects [56]. As demonstrated in Table 3, the results showed that all the model fit and quality indices are in the recommended range or have probability values less than .001. The APC is 0.237 for a $P < .001$, the ARS index is 0.529 for a $P < .001$, and the AARS index is 0.521 for a $P < .001$. All the values of APC, ARS, and AARS show a better fit than the recommended values. The AVIF is 2.246 and AFVIF is 2.324, representing there is no collinear problem found in the investigated model. The GoF is .649 which indicates a better fit than the large value of .36 [56]. In conclusion, the proposed model of mHealth is validated, representing good model fit and quality indices.

Table 3. Model fit and quality indices.

Quality indices	Statistics	Criteria (<i>P</i> value)	Result
Average path coefficient (APC)	0.237 (<i>P</i> <.001)	<.05	Fit
Average R-squared (ARS)	0.529 (<i>P</i> <.001)	<.05	Fit
Average adjusted R-squared (AARS)	0.521 (<i>P</i> <.001)	<.05	Fit
Average block variance inflation factor (AVIF)	2.246	Acceptable if ≤ 5 , ideally ≤ 3.3	Fit
Average full collinearity VIF (AFVIF)	2.324	Acceptable if ≤ 5.0 , ideally ≤ 3.3	Fit
Tenenhaus Goodness of Fit (GoF)	0.649	Small $\geq .1$, medium $\geq .25$, large $\geq .36$	Fit
R-squared contribution ratio (RSCR)	0.989	Acceptable if $\geq .9$, ideally =1.0	Fit

We further evaluated the psychometric properties of the instrument regarding reliability, convergent validity, and discriminate validity. According to the method used by Hair et al [58], Cronbach alpha and the composite reliability (CR) of each construct was used to test reliability and internal consistency. Table 4 showed that the values of Cronbach alpha and CR of all the constructs were higher than the recommended value (0.7) [56,58], thus exhibiting acceptable reliability and internal consistency. The validity of the measures was tested using convergent and discriminant validity. Fornell and Larcker [59] recommended that the average variance extracted (AVE) value should exceed .5 and each square correlation, which indicated adequate convergent validity and discriminant validity. As shown in Table 4, the AVE values of all constructs were between .686 and .898, which are greater than the recommended value (.5), thus demonstrating an excellent convergent validity. Furthermore, all the square roots of AVE were higher than any other correlation among the latent variables, thus indicating an adequate discriminant validity.

Structural Model

The structural research model was analyzed using WarpPLS 5.0 and the bootstrap resampling method [56]. Testing of the

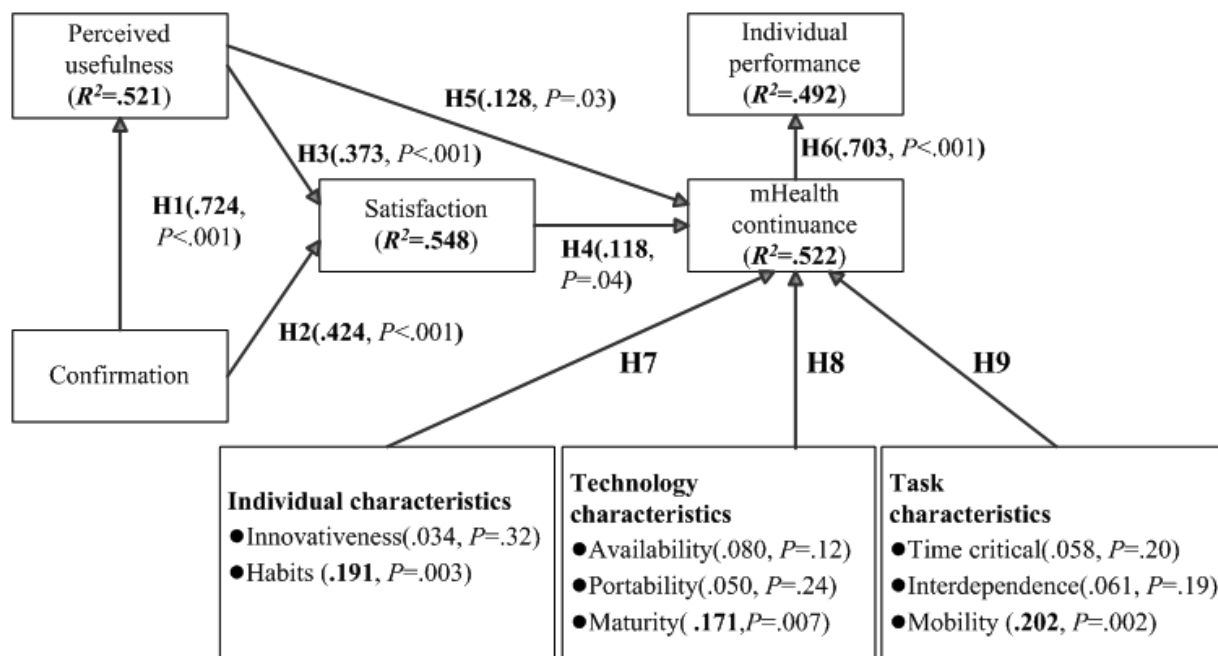
structural model was mainly on the basis of the path coefficient and R^2 value. Path coefficients represent the strength and direction of the relation among variables to test their significance, whereas R^2 values indicate the percentage to which external variables can explain the variability of internal variables and indicate the predictive power of the model. As shown in Figure 2, nine hypotheses (H1, H2, H3, H4, H5, H6, H7b, H8c, and H9c) were confirmed, whereas the remaining hypotheses (H7a, H8a, H8b, H9a, and H9b) were not significantly supported by this study. The results revealed that mHealth continuance ($R^2=0.522$) is mainly affected by perceived usefulness (beta=.128; *P*=.03), maturity (beta=.171; *P*=.007), habits (beta=.191; *P*=.003), task mobility (beta=.202; *P*=.002), and user satisfaction (beta=.118; *P*=.04), whereas individual performance ($R^2=0.492$) is affected by mHealth continuance use (beta=.703; *P*<.001). In addition, user satisfaction ($R^2=0.548$) is affected by confirmation (beta=.424; *P*<.001) and perceived usefulness (beta=.373; *P*<.001). Confirmation (beta=.724; *P*<.001) significantly affected perceived usefulness ($R^2=0.521$) and user satisfaction ($R^2=0.548$).

Table 4. Results of the reliability and validity of the research model.

Con-struct	CO ^a	PU ^b	SAT ^c	INN ^d	HAB ^e	AVA ^f	TC ^g	INT ^h	MC ⁱ	PER ^j	MOB ^k	PORT ^l	MAT ^m	AVE ⁿ (>.5)	CR ^o (>.7)	Cron- bach alpha (>.7)
CO	0.898	— ^p	—	—	—	—	—	—	—	—	—	—	—	0.806	0.943	.919
PU	0.718	0.869	—	—	—	—	—	—	—	—	—	—	—	0.755	0.939	.918
SAT	0.690	0.657	0.894	—	—	—	—	—	—	—	—	—	—	0.798	0.922	.874
INN	0.265	0.338	0.309	0.867	—	—	—	—	—	—	—	—	—	0.751	0.924	.889
HAB	0.586	0.500	0.502	0.320	0.910	—	—	—	—	—	—	—	—	0.829	0.951	.931
AVA	0.509	0.522	0.584	0.306	0.504	0.855	—	—	—	—	—	—	—	0.730	0.890	.815
TC	0.450	0.511	0.475	0.361	0.387	0.513	0.887	—	—	—	—	—	—	0.787	0.917	.864
INT	0.344	0.433	0.358	0.332	0.280	0.517	0.707	0.918	—	—	—	—	—	0.843	0.942	.906
MC	0.528	0.564	0.534	0.320	0.534	0.586	0.529	0.503	0.914	—	—	—	—	0.836	0.939	.901
PER	0.628	0.664	0.638	0.406	0.535	0.610	0.541	0.521	0.695	0.895	—	—	—	0.802	0.960	.950
MOB	0.302	0.406	0.286	0.245	0.326	0.410	0.476	0.582	0.463	0.444	0.947	—	—	0.897	0.946	.886
PORT	0.352	0.355	0.403	0.200	0.297	0.574	0.452	0.434	0.410	0.462	0.271	0.828	—	0.686	0.868	.771
MAT	0.440	0.477	0.506	0.249	0.340	0.653	0.547	0.517	0.518	0.606	0.337	0.621	0.899	0.809	0.927	.881

^aCO: confirmation.^bPU: perceived usefulness.^cSAT: satisfaction.^dINN: innovativeness.^eHAB: habits.^fAVA: availability.^gTC: time critical.^hINT: interdependence.ⁱMC: mobile health continuance.^jPER: performance.^kMOB: mobility.^lPORT: portability.^mMAT: maturity.ⁿAVE: average variance extracted.^oCR: composite reliability.^pThe omitted correlation coefficients between constructs in the upper diagonal matrix are equal to the values in lower diagonal matrix.

Figure 2. Results of the model validity. H1: The confirmation of mHealth systems significantly affects perceived usefulness; H2: The confirmation of mHealth systems significantly affects user satisfaction; H3: The perceived usefulness of mHealth systems significantly affects user satisfaction; H4: User satisfaction with mHealth systems significantly affects mHealth continuance; H5: The perceived usefulness of mHealth systems significantly affects mHealth continuance; H6: The continuance of mHealth significantly affects individual performance; H7: The individual characteristics of HCPs significantly affect mHealth continuance; H8: The technology characteristics of mHealth significantly affect mHealth continuance; H9: The task characteristics of HCPs significantly affect mHealth continuance; mHealth: mobile health.



Discussion

Key Factors Affecting Perceived Usefulness and User Satisfaction

Consistent with previous ECM-related studies [26,27], confirmation ($\beta=.724; P<.001$) is a factor that significantly affected perceived usefulness (by the HCPs) and user satisfaction. Both confirmation ($\beta=.424; P<.001$) and perceived usefulness ($\beta=.373; P<.001$) were significant predictors of user satisfaction with the mHealth systems. In the context of eHealth, confirmation accounted for 52.1% variance of perceived usefulness, whereas both confirmation and perceived usefulness accounted for 54.8% variance of satisfaction. Confirmation refers to users' perception of the congruence between the expectation of mHealth use and its actual performance [20], whereas perceived usefulness is the perception of users' regarding the expected benefits of mHealth use [20,41]. This implied that the expectations of the HCPs from the mHealth systems were confirmed through performance after implementation. The participants expected that mHealth use would positively affect the quality of clinical care, the effectiveness of medical teams for simultaneous processing of patient information, and patient care management. Understanding the expectations of the HCPs before system development and evaluating their responses after implementation of the mHealth system can increase the benefits of using the mHealth system in clinical care. Thus, continuous evaluation of whether the clinical care functions provided by the mHealth system meet the expectations of users is crucial. As the HCPs continuously use the system over a long time and become more

familiar with it, they may have new requirements for further improvement of the system. Dynamic changes in functional requirements should be considered by managers and system developers for ensuring user satisfaction. After mHealth is infused and integrated into the daily operations and clinical care practices of HCPs, the HCPs should compare their preadoption expectations and actual performance as well as perceived usefulness of the mHealth systems, thus improving the quality of clinical care and efficiency of care management. Such confirmations and perceived usefulness of mHealth are helpful in improving user satisfaction.

This study showed that perceived usefulness ($\beta=.128; P=.03$), user satisfaction ($\beta=.118; P=.04$), technology maturity ($\beta=.171; P=.007$), individual habits ($\beta=.191; P=.003$), and task mobility ($\beta=.202; P=.002$) exert significantly positive effects on mHealth continuance, which accounted for 52.2% of the total explained variance. Among the identified factors that affected mHealth continuance, task mobility, individual habits, and technology maturity have more significant direct effects on mHealth continuance than the factors (perceived usefulness and user satisfaction) derived from the ECM. Consistent with previous ECM-related studies [26,27], this study confirmed that perceived usefulness and user satisfaction were key predictors of mHealth continuance. In addition, the inclusion of characteristics of individuals, technology, and tasks facilitated the extension of the original ECM for understanding the factors influencing mHealth continuance. To increase the HCPs' intention toward mHealth continuance, paying attention to the characteristics related to task mobility, user habits, technology

maturity, and user perceptions related to perceived usefulness, and user satisfaction is necessary.

This study made an empirical validation on the framework of mHealth infusion proposed by O'Connor et al; however, only mobility, habit, and maturity were found to be salient predictors in mHealth infusion. Previous studies have highlighted that mobility is the primary reason for the applications of technological innovation in hospitals [14-16]. Zhang et al [49] reported that mHealth technologies offer the staff freedom to interact with and use technological tools irrespective of time and location. This study showed that mobility is the most crucial factor influencing mHealth continuance. Consistent with the results of a study by Limayem et al [22], this study revealed that habits play a major role in mHealth continuance. Previous studies have indicated that maturity is related to the existence of a level of system quality that is perceived as satisfactory and the perceived need for system improvement by the user [33,43,52]. O'Connor et al [33] argued that poor graphical user interface design and unsatisfactory process design of mobile systems result in unnecessary medical errors. When users (the HCPs) perceive poor quality of the mHealth systems, they are less likely to use the mHealth systems. Our study revealed that mHealth maturity was critical to its continuance. Therefore, it increased the intention of HCPs toward mHealth continuance by focusing on the design and implementation issues of the mHealth applications to satisfy actual users' needs. This study indicated that when the mHealth applications provide high quality (of mHealth system) and superior support for the HCPs' needs, the users (HCPs) had a relatively high intention toward mHealth continuance. As stated, we suggested the evaluation of task mobility, technology maturity, and individual habits and the provision of better support related to the fit among the aforementioned factors while introducing mHealth applications as those factors are salient predictors for mHealth continuance by the HCPs.

Furthermore, the perceived usefulness and user satisfaction of mHealth systems have been considered as critical factors affecting technology continuance in ECM-related studies [26,27], and they have been reported to exert significantly positive effects on mHealth continuance from the HCPs' perspective in this study. We should provide sufficient incentives and resources to improve perceived usefulness (by the HCPs) and user satisfaction after implementation of the system or in the infusion stage. Contrary to previous studies [33,34,49,50,53], we found that some factors in the characteristics of individuals (innovativeness), technology (availability and portability), and task (timeliness and interdependence) did not significantly affect mHealth continuance in this study. A possible explanation may be that aforementioned factors were not salient predictors in the mHealth context, particularly after implementation and in the infusion stage of the mHealth systems from the HCPs' perspective in Taiwan. In addition, those nonsignificant factors are mainly derived from the conceptual framework of mHealth infusion proposed by O'Connor et al [33] and they may obtain mixed results through empirical studies because of the difference

of research contexts, user groups, and application systems. It is acceptable that some factors investigated in this study were insignificant in the mHealth applications of the case hospital in Taiwan from the HCPs' perspective. For example, consumer' personal innovativeness is a significant factor of mHealth assimilation in Rai et al [34]; however, we found personal innovativeness is insignificant in the health care context from the HCPs' point of view.

Key Factors Affecting Individual Performance

The results indicated that mHealth continuance ($\beta=.703$; $P<.001$) exerted significantly positive effects on individual performance, thus explaining 49.2% variance in individual performance. O'Connor et al [33] found that the individual performance of HCPs was influenced by continued mHealth use in the infusion stage, which also significantly affected individual performance in system use. As expected, consistent with O'Connor et al [33] and Goodhue and Thompson [37], this study highlighted that mHealth continuance positively affected individual performance. If the HCPs intend to incorporate the mHealth systems into routine practices in the postimplementation or infusion stage, mHealth systems can enhance their individual performance, including improving the information exchange within a medical team and task identity in clinical care, increasing the efficiency of patient care, enhancing the quality of clinical patient care, and improving communication between health care personnel and patients or their families. System adoption in an organization is not always voluntary; sometimes it is because of work requirements or the necessity of IS for work completion. The ultimate goal of the system development process, including initial conception, implementation, adoption, and the following acceptance and continued use of mHealth, is to improve individual performance and satisfy clinical work demands. Therefore, we need to pay attention to system functions and demands of mHealth that require further improvement; thus, users will become more familiar and comfortable with mHealth. This is helpful to improve user work performance.

Furthermore, in this study, we evaluated the individual performance of the HCPs derived from mHealth continuance by using 6 items (Table 5). The results showed that the average score of each item ranged between 3.83 and 4.10, which indicated a positive evaluation by HCPs on mHealth continuance. According to the results in a descending ranking of the average score of each item, the HCPs perceived that the use of an eHealth system improved information exchange with the health care team (mean 4.10, SD 0.60), facilitated communication with patients and their families (mean 4.10, SD 0.60), provided efficient patient care (mean 3.94, SD 0.60), enhanced the quality of patient care (mean 3.91, SD 0.63), improved professional image (mean 3.86, SD 0.63), and facilitated task completion (mean 3.83, SD 0.62). This implied that improving information exchange with health care teams, facilitation of communication with patients and their families, and providing efficient patient care were the top 3 measures of performance of the mHealth systems identified by the HCPs.

Table 5. Individual performance derived from mobile health continuance.

Items	Mean (SD)
Using mHealth ^a can effectively improve information exchange between me and the health care team	4.10 (0.60)
Using mHealth can effectively facilitate my communication with patients and their families	4.10 (0.60)
Using mHealth allows me to provide efficient patient care	3.94 (0.60)
Using mHealth enhances the quality of patient care	3.91 (0.63)
Using mHealth improves my professional image	3.86 (0.63)
Using mHealth facilitates my work completeness	3.83 (0.62)
Average score	3.96 (0.61)

^amHealth: mobile health.

Conclusions

The key building block for sHealth care is mHealth, and the appropriate use of mHealth may result in major advances in expanding health care coverage, improving decision making, managing chronic conditions, and providing suitable health care during emergencies [9]. However, previous studies have indicated that mHealth is in its early stages of development [1]. Moreover, neither does current mHealth research adequately evaluate mHealth interventions nor does it provide sufficient evidence on the effects of mHealth on health [10]. Thus, appropriate evaluation, specifically after the implementation of mHealth systems and the use of the systems in daily health care practices, from the users' perspectives is critical. This study proposed an innovative extended model by integrating the ECM and characteristics of individuals, technology, and tasks to investigate critical factors affecting the continuance of mHealth and the performance of mHealth from the HCPs' perspective and assessing the infusion of the mHealth systems in clinical practices.

The results revealed that mHealth continuance was mainly affected by perceived usefulness, technology maturity, individual habits, task mobility, and user satisfaction, whereas individual performance was influenced by mHealth continuance. User satisfaction was affected by confirmation and perceived usefulness of mHealth, whereas perceived usefulness was affected by confirmation. This study showed that the ECM remained valid in the mHealth context from the HCPs' perspective. Among the identified factors that influenced mHealth continuance in this study, task mobility, individual habits, and technology maturity affected mHealth continuance more significantly than the factors (perceived usefulness and user satisfaction) derived from the ECM. To increase the intention of health professionals toward mHealth continuance, characteristics related to task mobility, user habits, and technology maturity and users' perceptions related to perceived usefulness and user satisfaction must be given attention.

We found that the users' intention toward mHealth continuance increased when the focus was on the design and implementation issues of the mHealth applications to satisfy the actual needs of users. This implied that if mHealth applications provided high quality of system and satisfactory support to meet the needs of the HCPs, the users will have a relatively high intention toward mHealth continuance. We further suggested the

evaluation of task mobility, technology maturity, and individual habits and provision of satisfactory support related to the fit between the aforementioned factors while introducing mHealth applications. Consistent with the results of previous ECM-related studies [26,27], we found that perceived usefulness and user satisfaction were the key factors affecting mHealth continuance from the HCPs' perspective. This study reported that confirmation played a key role in affecting perceived usefulness and user satisfaction. This indicated that the perceived usefulness and user satisfaction were effectively improved by minimizing the gaps between user expectations of mHealth use and its actual performance. We suggested the minimization of the gaps between user expectations of mHealth use and its actual performance by providing sufficient incentives and resources to improve the perceived usefulness and user satisfaction after implementation or in the infusion stage.

This study has made theoretical and practical contributions to the evaluation of mHealth systems. First, the study proposed an innovative integration model that extended the ECM with antecedents of IS infusion (including the characteristics of individuals, technology, and tasks) to identify the critical factors influencing mHealth continuance and performance from the perspective of HCPs. The extended ECM provided a comprehensive research model for investigating mHealth continuance or IS continuance. Second, the inclusion of characteristics of individual, technology, and task not only provided a reasonable framework but also highlighted that other studies can incorporate various critical factors depending on research contexts and situations. Third, the identified critical and salient factors that affected mHealth continuance and performance can be used as assessment tools by hospitals that have implemented mHealth to facilitate mHealth use and infusion. (4) The results can also help health care institutions that intend to introduce or develop mHealth applications in identifying critical issues and effectively allocating limited resources to mHealth systems.

We suggest focus areas for additional research and future studies on this topic. First, scholars can use the research model derived in this study, apply it to various research contexts, and compare the findings. Second, others can conduct an in-depth case study with the findings obtained from this study. To expand the research scope at the IS infusion stage, future studies should pay attention to the investigated factors (personal innovativeness, availability, portability, timeliness, and

interdependence) that were insignificant factors in this study. This is reasonable as those insignificant factors may have different (mixed) results because of the difference of research contexts, user groups, and application systems as mentioned in a summary of technology acceptance model studies [60].

This study has several limitations. First, this study was conducted only at a regional hospital in Taiwan; thus, the

findings obtained from this research may not be immediately transferrable to other countries with different participant demographics and cultures. Second, a cross-sectional survey design was used for this study; thus, the inherent limitations of the survey methodology were inevitable. Furthermore, this study sample comprised voluntary participants. However, as the survey approach is commonly used in the field, the use of this method may not have adversely affected the results.

Acknowledgments

The authors sincerely extend their gratitude and recognition to the National Science Council of Taiwan for funding this study (MOST 104-2410-H-041-007).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire for mobile health continuance and performance.

[[PDF File \(Adobe PDF File\), 42KB - medinform_v7i1e12350_app1.pdf](#)]

References

1. Solanas A, Patsakis C, Conti M, Vlachos IS, Ramos V, Falcone F, et al. Smart health: a context-aware health paradigm within smart cities. *IEEE Commun Mag* 2014 Aug;52(8):74-81. [doi: [10.1109/MCOM.2014.6871673](#)]
2. Varshney U. *Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring*. New York: Springer; 2009:288.
3. Divall P, Camosso-Stefinovic J, Baker R. The use of personal digital assistants in clinical decision making by health care professionals: a systematic review. *Health Informatics J* 2013 Mar;19(1):16-28. [doi: [10.1177/1460458212446761](#)] [Medline: [23486823](#)]
4. Mickan S, Tilson JK, Atherton H, Roberts NW, Heneghan C. Evidence of effectiveness of health care professionals using handheld computers: a scoping review of systematic reviews. *J Med Internet Res* 2013 Oct 28;15(10):e212 [FREE Full text] [doi: [10.2196/jmir.2530](#)] [Medline: [24165786](#)]
5. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](#)] [Medline: [22781312](#)]
6. van Velsen L, Beaujean DJ, van Gemert-Pijnen JE. Why mobile health app overload drives us crazy, and how to restore the sanity. *BMC Med Inform Decis Mak* 2013 Feb 11;13:23 [FREE Full text] [doi: [10.1186/1472-6947-13-23](#)] [Medline: [23399513](#)]
7. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P T* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](#)]
8. Wallace S, Clark M, White J. 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. *BMJ Open* 2012 Aug 24;2(4):e001099 [FREE Full text] [doi: [10.1136/bmjopen-2012-001099](#)] [Medline: [22923627](#)]
9. Varshney U. Mobile health: four emerging themes of research. *Decis Support Syst* 2014 Oct;66:20-35. [doi: [10.1016/j.dss.2014.06.001](#)]
10. Michael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Center for global health and economic development (CGHED) Earth Institute, Columbia University. New York; 2010 May. Barriers and gaps affecting mHealth in low and middle income countries. Policy white paper URL: http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf [accessed 2018-09-28] [WebCite Cache ID 72IHAmi9g]
11. WHO Global Observatory for eHealth. *mHealth: New Horizons for Health Through Mobile Technologies* 2011. Geneva, Switzerland: World Health Organization; 2011:102.
12. Avgerou C, Ciborra C, Land FF. *The Social Study Of Information and Communication Technology: Innovation, Actors, And Contexts*. Oxford and New York, NY: Oxford University Press; 2004.
13. Al-Ubaydli M. Handheld computers. *Br Med J* 2004 May 15;328(7449):1181-1184 [FREE Full text] [doi: [10.1136/bmj.328.7449.1181](#)] [Medline: [15142928](#)]
14. Wu IL, Li JY, Fu CY. The adoption of mobile healthcare by hospital's professionals: an integrative perspective. *Decis Support Syst* 2011 Jun;51(3):587-596. [doi: [10.1016/j.dss.2011.03.003](#)]
15. Lu YU, Xiao Y, Sears A, Jacko JA. A review and a framework of handheld computer adoption in healthcare. *Int J Med Inform* 2005 Jun;74(5):409-422. [doi: [10.1016/j.ijmedinf.2005.03.001](#)] [Medline: [15893264](#)]

16. Wu JH, Wang SC, Lin LM. Mobile computing acceptance factors in the healthcare industry: a structural equation model. *Int J Med Inform* 2007 Jan;76(1):66-77. [doi: [10.1016/j.jimedinf.2006.06.006](https://doi.org/10.1016/j.jimedinf.2006.06.006)] [Medline: [16901749](https://pubmed.ncbi.nlm.nih.gov/16901749/)]
17. Lin TC. Mobile nursing information system utilization: the task-technology fit perspective. *Comput Inform Nurs* 2014 Mar;32(3):129-137. [doi: [10.1097/CIN.0000000000000039](https://doi.org/10.1097/CIN.0000000000000039)] [Medline: [24419090](https://pubmed.ncbi.nlm.nih.gov/24419090/)]
18. Lu C, Hu Y, Xie J, Fu Q, Leigh I, Governor S, et al. The use of mobile health applications to improve patient experience: cross-sectional study in Chinese public hospitals. *JMIR Mhealth Uhealth* 2018 May 23;6(5):e126 [FREE Full text] [doi: [10.2196/mhealth.9145](https://doi.org/10.2196/mhealth.9145)] [Medline: [29792290](https://pubmed.ncbi.nlm.nih.gov/29792290/)]
19. Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the use of mobile health technology by patients: an observational study in primary care clinics. *JMIR Mhealth Uhealth* 2016 Apr 19;4(2):e41 [FREE Full text] [doi: [10.2196/mhealth.4928](https://doi.org/10.2196/mhealth.4928)] [Medline: [27095507](https://pubmed.ncbi.nlm.nih.gov/27095507/)]
20. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. *MIS Q* 2001 Sep;25(3):351-370. [doi: [10.2307/3250921](https://doi.org/10.2307/3250921)]
21. Oliver RL. A cognitive model for the antecedents and consequences of satisfaction. *J Marketing Res* 1980 Nov;17(4):460-469. [doi: [10.2307/3150499](https://doi.org/10.2307/3150499)]
22. Limayem M, Hirt SG, Cheung CM. How habit limits the predictive power of intention: The case of information systems continuance. *MIS Q* 2007 Dec;31(4):705-737. [doi: [10.2307/25148817](https://doi.org/10.2307/25148817)]
23. Lin CS, Wu S, Tsai RJ. Integrating perceived playfulness into expectation-confirmation model for web portal context. *Inform Manage* 2005 Jul;42(5):683-693. [doi: [10.1016/j.im.2004.04.003](https://doi.org/10.1016/j.im.2004.04.003)]
24. Chou SW, Min HT, Chang YC, Lin CT. Understanding continuance intention of knowledge creation using extended expectation-confirmation theory: an empirical study of Taiwan and China online communities. *Behav Inf Technol* 2010 Nov;29(6):557-570. [doi: [10.1080/01449290903401986](https://doi.org/10.1080/01449290903401986)]
25. Mettler T. Post-acceptance of electronic medical records: evidence from a longitudinal field study. 2012 Dec 14 Presented at: Proceedings of the 33rd International Conference on Information Systems; December 16-19, 2012; Orlando, Florida, USA URL: <https://tinyurl.com/yxoxsb5d>
26. Akter S, Ray P, D'Ambra J. Continuance of mHealth services at the bottom of the pyramid: the roles of service quality and trust. *Electron Markets* 2012 May 8;23(1):29-47. [doi: [10.1007/s12525-012-0091-5](https://doi.org/10.1007/s12525-012-0091-5)]
27. Bhattacherjee A. Acceptance of e-commerce services: the case of electronic brokerages. *IEEE Trans. Syst., Man, Cybern. A* 2000;30(4):411-420. [doi: [10.1109/3468.852435](https://doi.org/10.1109/3468.852435)]
28. Chen SC, Chen HH, Chen MF. Determinants of satisfaction and continuance intention towards self - service technologies. *Ind Manage Data Syst* 2009 Oct 30;109(9):1248-1263. [doi: [10.1108/02635570911002306](https://doi.org/10.1108/02635570911002306)]
29. Chen SC, Jong D, Lai MT. Assessing the relationship between technology readiness and continuance intention in an E-appointment system: relationship quality as a mediator. *J Med Syst* 2014 Sep;38(9):76. [doi: [10.1007/s10916-014-0076-3](https://doi.org/10.1007/s10916-014-0076-3)] [Medline: [25007955](https://pubmed.ncbi.nlm.nih.gov/25007955/)]
30. Parthasarathy M, Bhattacherjee A. Understanding post-adoption behavior in the context of online services. *Inform Syst Res* 1998;9(4):362-379.
31. Bhattacherjee A, Barfar A. Information technology continuance research: current status and future directions. *Asia Pac J Inf Syst* 2011;21(2):1-18 [FREE Full text]
32. Hsiao JL, Chen RF. An investigation on task-technology fit of mobile nursing information systems for nursing performance. *Comput Inform Nurs* 2012 May;30(5):265-273. [doi: [10.1097/NCN.0b013e31823eb82c](https://doi.org/10.1097/NCN.0b013e31823eb82c)] [Medline: [22156768](https://pubmed.ncbi.nlm.nih.gov/22156768/)]
33. O'Connor Y, O'Reilly P, O'Donoghue J. M-health infusion by healthcare practitioners in the national health services (NHS). *Health Policy Techn* 2013 Mar;2(1):26-35 [FREE Full text] [doi: [10.1016/j.hlpt.2012.12.002](https://doi.org/10.1016/j.hlpt.2012.12.002)]
34. Rai A, Chen L, Pye J, Baird A. Understanding determinants of consumer mobile health usage intentions, assimilation, and channel preferences. *J Med Internet Res* 2013 Aug 2;15(8):e149 [FREE Full text] [doi: [10.2196/jmir.2635](https://doi.org/10.2196/jmir.2635)] [Medline: [23912839](https://pubmed.ncbi.nlm.nih.gov/23912839/)]
35. DeLone WH, McLean ER. Information systems success: the quest for the dependent variable. *Inform Syst Res* 1992 Mar;3(1):60-95. [doi: [10.1287/isre.3.1.60](https://doi.org/10.1287/isre.3.1.60)]
36. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. *J Manage Inform Syst* 2003;19(4):9-30. [doi: [10.1080/07421222.2003.11045748](https://doi.org/10.1080/07421222.2003.11045748)]
37. Goodhue DL, Thompson RL. Task-technology fit and individual performance. *MIS Q* 1995 Jun;19(2):213-236. [doi: [10.2307/249689](https://doi.org/10.2307/249689)]
38. Tennant V, Mills A, Chin W. Investigating information system infusion at the individual level: Re-conceptualisation and operationalization. : Proceedings of the 15th Pacific Asia Conference on Information Systems 2011; 2011 Presented at: 15th Pacific Asia Conference on Information Systems; July 7-11, 2011; Brisbane, Australia p. 189.
39. Hsiao JL, Chang HC, Chen RF. A study of factors affecting acceptance of hospital information systems: a nursing perspective. *J Nurs Res* 2011 Jun;19(2):150-160. [doi: [10.1097/JNR.0b013e31821cbb25](https://doi.org/10.1097/JNR.0b013e31821cbb25)] [Medline: [21586992](https://pubmed.ncbi.nlm.nih.gov/21586992/)]
40. Chen RF, Hsiao JL. An investigation on physicians' acceptance of hospital information systems: a case study. *Int J Med Inform* 2012 Dec;81(12):810-820. [doi: [10.1016/j.jimedinf.2012.05.003](https://doi.org/10.1016/j.jimedinf.2012.05.003)] [Medline: [22652011](https://pubmed.ncbi.nlm.nih.gov/22652011/)]

41. Kuo KM, Liu CF, Ma CC. An investigation of the effect of nurses' technology readiness on the acceptance of mobile electronic medical record systems. *BMC Med Inform Decis Mak* 2013 Aug 12;13:88 [FREE Full text] [doi: [10.1186/1472-6947-13-88](https://doi.org/10.1186/1472-6947-13-88)] [Medline: [23938040](https://pubmed.ncbi.nlm.nih.gov/23938040/)]
42. Wright E, Marvel J. Electronic health records: postadoption physician satisfaction and continued use. *Health Care Manag (Frederick)* 2012;31(3):259-267. [doi: [10.1097/HCM.0b013e3182619e90](https://doi.org/10.1097/HCM.0b013e3182619e90)] [Medline: [22842761](https://pubmed.ncbi.nlm.nih.gov/22842761/)]
43. Gebauer J, Shaw MJ, Gribbins ML. Task-technology fit for mobile information systems. *J Inf Technol* 2010 Sep;25(3):259-272. [doi: [10.1057/jit.2010.10](https://doi.org/10.1057/jit.2010.10)]
44. Junglas I, Abraham C, Ives B. Mobile technology at the frontlines of patient care: understanding fit and human drives in utilization decisions and performance. *Decis Support Syst* 2009 Feb;46(3):634-647. [doi: [10.1016/j.dss.2008.11.012](https://doi.org/10.1016/j.dss.2008.11.012)]
45. Agarwal R, Prasad J. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Inform Syst Res* 1998 Jun;9(2):204-215. [doi: [10.1287/isre.9.2.204](https://doi.org/10.1287/isre.9.2.204)]
46. Sun H. Understanding User Revisions When Using Information System Features: Adaptive System Use and Triggers. *MIS Quarterly* 2012 Jun;36(2):453-478. [doi: [10.2307/41703463](https://doi.org/10.2307/41703463)]
47. Ouellette JA, Wood W. Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychol Bull* 1998;124(1):54-74. [doi: [10.1037/0033-2909.124.1.54](https://doi.org/10.1037/0033-2909.124.1.54)]
48. Gefen D. TAM or just plain habit: a look at experienced online shoppers. *J Organ End User Comp* 2003;15(3):1-13. [doi: [10.4018/joeuc.2003070101](https://doi.org/10.4018/joeuc.2003070101)]
49. Zhang N, Guo X, Wang F, Chen G, Wei Q. Task-technology fit in mobile work: Exploring the links between task attributes and technology characteristics. 2011 Jun Presented at: 10th International Conference on Mobile Business; June 20-21, 2011; Como, Italy. [doi: [10.1109/ICMB.2011.47](https://doi.org/10.1109/ICMB.2011.47)]
50. Liden RC, Wayne S, Bradway LK. Task interdependence as a moderator of the relation between group control and performance. *Hum Relat* 1997 Feb;50(2):169-181. [doi: [10.1023/A:1016921920501](https://doi.org/10.1023/A:1016921920501)]
51. O'Connor Y, O'Raghailigh PJ, O'Donoghue J. Individual infusion of m-health technologies: Determinants and outcomes. 2012 Jun Presented at: ECIS 2012; June 10-13, 2012; Barcelona, Spain.
52. Triandis HC. Values, attitudes, and interpersonal behavior. *Nebr Symp Motiv* 1980;27:195-259. [Medline: [7242748](https://pubmed.ncbi.nlm.nih.gov/7242748/)]
53. Hoehle H, Scornavacca E. Unveiling experts perceptions towards the characteristics and value propositions of mobile information systems. : *IEEE*; 2008 Jul Presented at: 7th International Conference on Mobile Business; July 7-8, 2008; Barcelona, Spain. [doi: [10.1109/ICMB.2008.8](https://doi.org/10.1109/ICMB.2008.8)]
54. Petrick JF. Development of a multi-dimensional scale for measuring the perceived value of a service. *J Leisure Res* 2017 Dec 13;34(2):119-134. [doi: [10.1080/00222216.2002.11949965](https://doi.org/10.1080/00222216.2002.11949965)]
55. Lowry PB, Gaskin J. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Trans Prof Commun* 2014 Jun;57(2):123-146. [doi: [10.1109/TPC.2014.2312452](https://doi.org/10.1109/TPC.2014.2312452)]
56. Kock N. ScriptWarp Systems. Laredo, Texas: ScriptWarp Systems; 2015. WarpPLS 5.0 User Manual URL: http://cits.tamui.edu/WarpPLS/UserManual_v_5_0.pdf [accessed 2018-09-28] [WebCite Cache ID 72JJO1Y2N]
57. Chin WW. Bootstrap cross-validation indices for PLS path model assessment. In: *Handbook Of Partial Least Squares: Concepts, Methods And Applications* (springer Handbooks Of Computational Statistics). Berlin, Heidelberg: Springer; 2010:83-97.
58. Hair Jr J, Sarstedt M, Hopkins L, Kuppelwieser VG. Partial least squares structural equation modeling (PLS-SEM): an emerging tool in business research. *Eur Bus Rev* 2014 Mar 4;26(2):106-121. [doi: [10.1108/EBR-10-2013-0128](https://doi.org/10.1108/EBR-10-2013-0128)]
59. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 1981 Feb;18(1):39-50. [doi: [10.2307/3151312](https://doi.org/10.2307/3151312)]
60. Lee Y, Kozar KA, Larsen KR. The technology acceptance model: past, present, and future. *commun assoc inform syst* 2003;12(1):752-780. [doi: [10.17705/1CAIS.01250](https://doi.org/10.17705/1CAIS.01250)]

Abbreviations

- AARS:** average adjusted R-squared
- AFVIF:** average full collinearity variance inflation factor
- APC:** average path coefficient
- ARS:** average R-squared
- AVE:** average variance extracted
- AVIF:** average block variance inflation factor
- CR:** composite reliability
- CVI:** content validity index
- ECM:** expectation-confirmation model
- eHealth:** electronic health
- GoF:** Goodness of Fit
- HCP:** health care professional

HIS: hospital information system
ICT: information communication technology
IS: information system
IT: information technology
mHealth: mobile health
PC: personal computer
PLS: partial least square
RSCR: R-squared contribution ratio
sHealth: smart health
TTF: task-technology fit

Edited by G Eysenbach; submitted 28.09.18; peer-reviewed by A Baird, KM Kuo; comments to author 25.10.18; revised version received 16.12.18; accepted 31.12.18; published 18.03.19.

Please cite as:

Hsiao JL, Chen RF

Understanding Determinants of Health Care Professionals' Perspectives on Mobile Health Continuance and Performance

JMIR Med Inform 2019;7(1):e12350

URL: <http://medinform.jmir.org/2019/1/e12350/>

doi: [10.2196/12350](https://doi.org/10.2196/12350)

PMID: [30882353](https://pubmed.ncbi.nlm.nih.gov/30882353/)

©Ju-Ling Hsiao, Rai-Fu Chen. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 18.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

European Hospitals' Transition Toward Fully Electronic-Based Systems: Do Information Technology Security and Privacy Practices Follow?

Sylvestre Uwizeyemungu^{1*}, DBA; Placide Poba-Nzaou^{2*}, DBA; Michael Cantinotti^{3*}, PhD

¹Accounting Department, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

²Department of Organization and Human Resources Management, École des Sciences de la Gestion, Université du Québec à Montréal, Montréal, QC, Canada

³Psychology Department, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

* all authors contributed equally

Corresponding Author:

Sylvestre Uwizeyemungu, DBA

Accounting Department

Université du Québec à Trois-Rivières

3351 Boul. des Forges

Trois-Rivières, QC, G9A 5H7

Canada

Phone: 1 819 376 5011 ext 3164

Fax: 1 819 376 5180

Email: sylvestre.uwizeyemungu@uqtr.ca

Abstract

Background: Traditionally, health information has been mainly kept in paper-based records. This has deeply changed throughout approximately the last three decades with the widespread use of multiple health information technologies. The digitization of health care systems contributes to improving health care delivery. However, it also exposes health records to security and privacy breaches inherently related to information technology (IT). Thus, health care organizations willing to leverage IT for improved health care delivery need to put in place IT security and privacy measures consistent with their use of IT resources.

Objective: In this study, 2 main objectives are pursued: (1) to assess the state of the implementation of IT security and privacy practices in European hospitals and (2) to assess to what extent these hospitals enhance their IT security and privacy practices as they move from paper-based systems toward fully electronic-based systems.

Methods: Drawing on data from the European Commission electronic health survey, we performed a cluster analysis based on IT security and privacy practices implemented in 1723 European hospitals. We also developed an IT security index, a compounded measure of implemented IT security and privacy practices, and compared it with the hospitals' level in their transition from a paper-based system toward a fully electronic-based system.

Results: A total of 3 clearly distinct patterns of health IT-related security and privacy practices were unveiled. These patterns, as well as the IT security index, indicate that most of the sampled hospitals (70.2%) failed to implement basic security and privacy measures consistent with their digitization level.

Conclusions: Even though, on average, the most electronically advanced hospitals display a higher IT security index than hospitals where the paper system still dominates, surprisingly, it appears that the enhancement of IT security and privacy practices as the health information digitization advances in European hospitals is neither systematic nor strong enough regarding the IT-security requirements. This study will contribute to raising awareness among hospitals' managers as to the importance of enhancing their IT security and privacy measures so that they can keep up with the security threats inherently related to the digitization of health care organizations.

(*JMIR Med Inform* 2019;7(1):e11211) doi:[10.2196/11211](https://doi.org/10.2196/11211)

KEYWORDS

health information technology; data security; patient data privacy; health services; electronic health records

Introduction

Motivation and Objectives

In many countries, health care services delivery is being reformed—and some say revolutionized [1]—through information technology (IT). IT potential is being leveraged in the quest to achieve what has been called the *triple aim* [2-4], that is, (1) improving individual care experience, (2) improving population health, and (3) reducing per capita cost of health care.

IT developments have led to the digitization of health records, thus offering new or improved means to efficiently and effectively collect, process, store, consult, and share health information. Digitized, health information becomes more portable and readily shareable within and among different health care organizations; it becomes readily available to public health administrators for health surveillance and policy-making purposes; it becomes, under certain conditions, available for research; it also becomes more accessible to patients. Thus far, the significant majority of the literature suggests positive effects of digitization on the effectiveness of health care outcomes [5]. However, despite all those advantages, the digitization of health information exposes health records to security breaches inherently related to IT [6]. Indeed, potential users of health IT (HIT) express various concerns over IT-related security and privacy issues [7], and these concerns may negatively affect the trust of potential HIT users. As a result of this decrease in trust, patients, as well as health care professionals, may be reluctant to widely use some HIT functionalities, such as health information exchange (HIE), telehealth, and mobile health [8], threatening the necessary “meaningful use” of HIT [9]. The ultimate result would be ineffective health care delivery [10,11], as well as ineffective public health monitoring [12] or health research [13,14].

To alleviate security and privacy concerns, health care organizations willing to leverage IT for improved health care delivery need to put in place IT security measures consistent with their IT development plans. In this study, our objective was twofold: to assess the state of the implementation of IT security and privacy practices in European hospitals and to assess to what extent European hospitals enhance their IT security practices as they move from paper-based systems toward fully electronic-based systems.

Background and Significance

Traditionally, health information has been mainly kept in paper-based records. This has deeply changed throughout approximately the last three decades, with the widespread use of multiple HIT, an umbrella term we use here to refer to all IT systems used for storing, accessing, processing, sharing, transmitting health information, or for supporting health care delivery and health care systems management. Thus defined, HIT encompasses all the 4 functionality-based categories of IT proposed by Adler-Milstein et al [15]: provider-centric electronic record, patient-centric electronic record, HIE, and telehealth.

By their mere nature, HIT compiles a wide range of highly sensitive information. This information includes not only current

data related to patients’ tests, diagnoses, and treatments but also past medical history [16]. Health care providers need to keep and manipulate that information securely not only to meet the patients’ willingness to keep their health information private but also to live up to the health care organizations’ moral and legal responsibilities. However, the task of keeping health records secure is affected by the dynamic nature of the HIT environment.

In recent years, the landscape of security of health records has changed following a number of phenomena related to the digitization. For example, professional and academic literature has extensively echoed security issues arising from IT trends such as hosting health records on distant servers operated by third-party cloud services providers [17,18], the usage of mobile devices, and the related trends of bring your own device in health care [19-21], as well as mobile health apps [22,23] and the IT-enabled HIE [24,25].

Information Technology Security Incidents in Health Care Settings

A report by Infosec Institute underscored that the remarkable growth in the adoption of electronic health records (EHRs) in the last years has not been accompanied by a parallel evolution in cybersecurity measures, thus rendering the health care industry ill-equipped and poorly protected with regard to cyber threats [26]. This bleak assessment seems to be upheld by numerous reports of IT-related incidents in hospital settings. According to a 2014 survey by Information Security Media Group (ISMG), at least one security breach that affects fewer than 500 individuals has occurred in 75% of surveyed health care organizations in the United States, and at least one incident affecting more than 500 individuals has been reported by 21% of surveyed health care providers [27]. In the 2015 survey by Healthcare Information and Management Systems Society, two-thirds (68%) of surveyed health care organizations in the United States reported having recently experienced a significant security incident [28]. Reported security incidents came both from external threats (63.6% of health care organizations) and insider threats (53.7%) [28].

These statistics of IT-related security breaches in health care settings are disturbing, and the reality may be even bleaker when one considers that many security incidents remain undetected or are not properly assessed [27], as well as the propensity of organizations to underreport security incidents [29].

Documented incidents show that security breaches in health care settings can be expensive. For example, Absolute Software Corporation reported cases of breaches in health care data that cost hospitals from US \$250,000 to US \$2.5 million in settlement payments. Even though these amounts are quite sizeable, they represent but a fraction of the overall financial burden of the incidents [30].

Security and privacy concerns, as well as the fear of related liabilities, may prevent health care providers from leveraging IT for improving their services. Increasing HIT security and privacy practices in hospitals is then an important step forward for effective health care delivery.

Health Information Technology–Related Security System

In response to IT-related security and privacy concerns, health care providers who adopt HIT need to put in place an adequate security system. This system is “a set of security mechanisms that are implemented according to a security policy,” which is “a collection of rules that allow or disallow possible actions, events, or something related to security” [10].

Generally speaking, an IT security policy aims at ensuring that an organization’s IT assets (hardware, software, data, and people) respond constantly to required levels of confidentiality, integrity, and availability [31,32]. These 3 basic IT security requirements are generally referred to as the CIA triad (Confidentiality, Integrity, Availability) [33].

The notion of confidentiality is generally defined as “restricting information to persons belonging to a set of specifically authorized recipients [34].” Confidentiality requires that only duly authorized people can get access to data, whether they are stored, being transmitted, or being treated. This can be achieved through encryption of data or through controlled access to the systems. These are technological means, but confidentiality can also be achieved relying on moral dispositions (eg, professional silence) [34]. With regard to encryption, the 2014 survey of ISMG showed that although encryption is commonly applied for health data transmitted across exposed networks, it is less applied to data stored in mobile devices and other storage media [27]. The confidentiality requirement responds to privacy concerns that are of paramount importance in health care systems given the sensitivity of information they contain.

With the integrity criterion, it is expected “that information is protected against unauthorized modification or deletion as well as irrevocable, accidental, and undesired changes by authorized users” [33].

As for availability, it requires that a system be accessible and fully operational whenever an authorized user needs to utilize it. The availability criterion refers to multiple aspects ranging from scalability (adaptability to changing performance needs) to resilience (resistance to software or hardware failures) and to recoverability of data in case of loss for whatever reason [33].

Methods

Data Source

We used data from the European Commission 2013 electronic health (eHealth) survey (Joint Research Centre, Institute for Prospective Technological Studies). The objective of the survey was “to benchmark the level of eHealth use in acute care hospitals in all 27 European Union member states, Croatia, Iceland, and Norway” [35]. The dataset used might seem outdated but we deem it still relevant. First, the survey we are referring to is the last one of this magnitude to have been conducted at the level of the European Union. It is thus the most recent. Second, it has been demonstrated that secondary data that are 5 years old or even older can provide meaningful, empirically grounded, and useful insights in the field of IT security in the health care sector [36]. Third, over the last 5

years, the HIT security field has not recorded significant changes that would render our data and our analysis obsolete. The analysis of recent literature reviews of IT security in health care settings [37–39] shows that the health care industry still lags behind in IT security measures implementation. It also shows that there is no major technical breakthrough in HIT security and no notable new-brand threats. One can only note a greater awareness among health care professionals and patients following multiple breaches of health data made public and the strengthening of legal requirements. In this regard, the adoption of the General Data Protection Regulation throughout the European Union territory has been too recent for its effects on IT security and privacy practices in the health sector to have been felt. This regulation came into force in May 2018 [40]. Fourth, other factors specific to the hospital context suggest a slow pace of changes in such a context. Many small health care organizations lack financial and human resources for undertaking substantive (as opposed to symbolic) IT security programs [41]. For their part, large health care organizations tend to be complex systems in their structure and management [42,43], and in such systems, cultural shifts to implement security and privacy measures may take a lot of time.

Sample

To ensure the representativeness of the sample, the survey team undertook the following steps [44]. They estimated the universe of acute care hospitals in the European Union, combining various sources (previous survey, lists of hospitals from the World Health Organization, and national ministry of health of each covered country). The estimation yielded a universe of 8199 acute care hospitals. From this point, the survey team proceeded to a stratification sampling procedure to ensure geographical representation (Nordic countries, Southern Europe, Western Europe, and Central and Eastern Europe). The stratification process also included other considerations related to the hospital’s ownership (public, private, and other) and size (number of beds). To guarantee the representativeness of the sample at the end of data collection, the survey team proceeded to nonresponse rate corrections. In total, the survey team contacted 5424 acute care hospitals, and 1753 hospitals completed the interview [45]. This corresponds to a rate response of 32.3%.

Of the 1753 initial observations, only 30 (1.7%) were dropped because of missing data (“don’t know” response or no answer at all) on key variables, which led us to a final sample of 1723 European hospitals. Even though the portion of dropped cases is very low, we performed statistical analyses to assess whether the dropped observations were significantly different in any way from the retained observations. Little’s missing completely at random test ($\chi^2_{5,1723}=7.8, P=.17$) indicated that ignoring cases with missing values does not a priori bring about a systematic bias. However, we went a step further and performed a nonresponse bias analysis comparing key characteristics of dropped cases against sampled cases (Table 1). To do so, we used Fisher exact test of homogeneity instead of chi-square test, thus taking into account the fact that some cells of the contingency table would contain low expected frequencies (<5), especially for the group of dropped cases (n=30). The results

showed that the dropped cases are not significantly different from the sampled cases with regard to all of hospital characteristics.

Table 1 depicts descriptive statistics of our sample. Public hospitals account for 70.38% (1188/1688) of surveyed hospitals. Most hospitals in our sample are not university hospitals (1485/1723; 86.19%). Independent hospitals operating either on 1 site (710/1723; 41.21%) or on multiple sites (542/1723; 31.46%) make up 72.67% (1252/1723) of the sample. More than half (857/1568; 54.66%) of surveyed hospitals can be qualified as small (250 or fewer beds), whereas a small portion

(167/1568; 10.65%) falls in the category of large hospitals (more than 750 beds). With regard to IT budget, a large majority (1086/1260; 86.19%) of sampled hospitals allocate 3% or less of their total budget to their IT function. Hospitals that devote more than 5% of their total budget to IT represent a negligible portion of the sample (52/1260; 4.13%). Almost 7 out of 10 hospitals (1159/1665; 69.61%) have in place an in-house, designed IT security regulation; 6 out of 10 (997/1665; 59.88%) report relying on a nationally-designed regulation, whereas a regional regulation is referred to by less than 3 out of 10 hospitals (475/1665; 28.53%).

Table 1. Characteristics of sampled versus nonsampled hospitals.

Variable and characteristics	% of nonsampled (n=variable), n (%)	% of sampled (n=variable), n (%)	Chi-square (df)	P value
Status	N=28	N=1688	0.2 (2)	.91
Public	20 (71)	1188 (70.38)		
Private	5 (18)	335 (19.85)		
Not for Profit	3 (11)	165 (9.77)		
University hospital	N=30	N=1723	1.3 (1)	.42
Yes	2 (7)	238 (13.81)		
No	28 (93)	1485 (86.19)		
Single/multiple sites	N=30	N=1723	1.7 (4)	.77
Independent/One site	14 (47)	710 (41.21)		
Independent/Multiple sites	9 (30)	542 (31.46)		
Part of a group of hospitals	4 (13)	341 (19.79)		
Part of a group of care institutions	2 (7)	78 (4.53)		
Other	1 (3)	52 (3.02)		
Size (number of beds)	N=25	N=1568	7.4 (3)	.05
<101	11 (44)	363 (23.15)		
101 to 250	8 (32)	494 (31.51)		
251 to 750	6 (24)	544 (34.69)		
>750	0 (0)	167 (10.65)		
IT^a budget (% of total hospital budget)	N=19	N=1260	1.6 (3)	.62
<1%	10 (53)	486 (38.57)		
1% to 3%	7 (37)	600 (47.62)		
3.1% to 5%	2 (11)	122 (9.68)		
>5%	0 (0)	52 (4.13)		
Security regulation	N=23	N=1665		
National	11 (48)	997 (59.88)	1.4 (1)	.29
Regional	3 (13)	475 (28.53)	2.7 (1)	.16
Hospital	13 (57)	1159 (69.61)	1.8 (1)	.18

^aIT: information technology.

Table 2. Security and privacy practices measures.

Variable	Measure (Yes or no)
Confidentiality: which of the following security measures are taken to protect the patient data stored and transmitted by the hospital's IT^a system?	
Stored data	Encryption of stored data
Transmitted data	Encryption of transmitted data
Access control	Workstations with access only through health professional cards or codes
Integrity	Is data entry in the hospital's IT system certified with digital signature?
Availability	Is your IT team able to immediately restore critical clinical information system operations if a disaster causes the complete loss of data at your hospital's primary data center?

^aIT: information technology.

Measurement

For contextual variables, measures used are depicted in [Table 1](#), column “characteristics.” The level attained by a hospital in the transition from a paper-based system to a fully electronic-based system was measured on a 9-point Likert scale (1=totally paper-based and 9=totally electronic-based, with point 5 as a hybrid model). For simplicity, this measure has been transformed into 3 nominal categories: hospitals that chose positions from 1 to 3 were qualified as having a paper-dominant system; a hybrid system label was given to hospitals that chose positions from 4 to 6, and the remaining hospitals (positions from 7 to 9) were deemed to have an electronic-dominant system. We present in [Table 2](#) the measures used to capture the implementation of HIT security and privacy practices used as our clustering variables. These practices were measured through a dichotomous scale (eg, coding “1” if a confidentiality-related practice is implemented and “0” when it is not implemented).

To check for possible variables multicollinearity before performing our cluster analysis [46], we produced a correlation matrix of all our variables (clustering variables and contextual variables; [Multimedia Appendix 1](#)). The results show that the risk of multicollinearity is very low: all the correlations coefficients are very low, the maximum correlation coefficient noted being .359 (between encryption of stored data and encryption of transmitted data). Besides, multicollinearity diagnostics ([Multimedia Appendix 1](#)) show that the tolerance values vary between .78 and .96, far above the commonly used cutoff threshold of .1 [47]. The highest variance inflation factor noted is equal to 1.28, far below the usual cutoff threshold of 10.0 [47]. These values show that multicollinearity is not a concern in our data.

To assess the state of the implementation of IT security and privacy practices in European hospitals, we developed an IT security index for each hospital. We developed the IT security index based on the responses given by hospitals with regard to their security practices. We considered 5 questions asked (see [Table 2](#)). A given hospital had to respond whether or not it has implemented the security practice related to confidentiality (3 questions), to integrity (1 question), and to availability (1 question). With a “Yes” response, a hospital scored 1 point and 0 otherwise (a “No” answer). The IT security index reflects the total score obtained by each hospital for the 5 questions. However, considering that the confidentiality component was

measured through 3 questions, whereas the other 2 components (integrity and availability) were measured each by 1 question, we estimated that the score obtained by simply summing up the results from the 5 questions would be skewed in favor of the confidentiality component. In the absence of theoretical or empirical evidence to the effect that any of the 3 components of the security triad (confidentiality-integrity-availability) would be more significant when compared with the 2 others, we hypothesized an equal weight for the 3 components. That is why the confidentiality component accounts for 1 in the index even if it is measured through 3 questions: each of the 3 questions related to confidentiality accounted for one-third (1/3) of the value of confidentiality in the index. As the integrity and availability components account each for 1, the maximum score of the IT security index for a given hospital is 3 (for a hospital that has all the 5 practices implemented: $1/3+1/3+1/3+1+1$), and the minimum is 0 (for a hospital with none of the practices implemented).

Cluster Analysis

We performed an agglomerative hierarchical clustering procedure combining Ward's minimum variance criterion with the squared Euclidian distance. This procedure allows distributing the observations into distinct subgroups (clusters) in the way that maximizes at the same time the intrasubgroup similarity and the intersubgroups dissimilarity [48]. Each subgroup comprises hospitals more or less homogenous with regard to clustering criteria (in this case HIT security practices implemented), and each subgroup is highly distinct to other subgroups with regard to the same criteria.

To identify the optimal number of clusters, we first examined the Euclidian distances across the clusters in the dendrogram produced with the clustering procedure. We identified 2 apparently equally plausible solutions, a 3-cluster and a 4-cluster solution. To decide which of these 2 solutions would be better, we followed Ketchen and Shook's [49] recommendation: we ascertained the robustness of both by replicating the clustering algorithm on subsamples of about 80%, 60%, and 40% of observations randomly selected using SPSS's random selection functionality (SPSS version 24, SPSS Inc). The analysis of the dendrograms produced with all these subsamples suggested a slight advantage of the 4-cluster solution over the 3-cluster solution. However, when we performed a discriminant analysis test, we found that 2 clusters of the 4-solution clusters were too

close, whereas all clusters in the 3-cluster solution were well-separated. Therefore, we chose the 3-cluster solution.

The discriminant analysis also allowed us to test the validity of the clusters. This test “runs the data back through the minimum-variance method as a discriminant function to see how accurately hospitals are classified [50].” The results of this test indicated a perfect classification accuracy (100%) for clusters 2 and 3 and a high level of classification accuracy (92.3%) for cluster 1. Overall, 98.7% of original observations were correctly classified. Moreover, based on cross-validation with analysis and holdout subsamples of respectively 60% (n=992) and 40% (n=731) of the total sample [51], for the 3-cluster solution, both hit ratios (97.7% for the analysis subsample and 98.1% for the holdout subsample) largely exceeded the threshold values of both maximum chance criteria (C_{max}) and proportional chance criteria (C_{Pro}). Indeed, the 2 hit ratios should be greater than $Max(1.25[C_{Pro}; C_{max}])$, which is the case as in this study $C_{Pro}=34\%$ and $C_{max}=39.3\%$. Thus, the null hypothesis that the percentage correctly classified was not significantly different from what would be classified by chance alone was rejected.

Results

Implementation Level of Information Technology Security and Privacy Practices

For the whole sample, the mean IT security index is 1.26 (with an SD of 0.83), the median being 1.33. In Table 3, we present the detailed statistics of hospitals by IT security index level.

It can be noted that only a tiny fraction of hospitals (50/1723; 2.90%) display a perfect IT security index (3). The nonnegligible fraction of hospitals (225/1723; 13.06%) has no IT security and privacy practices implemented whatsoever (IT security index=0). From the column “cumulative %,” one can note the total percentage of hospitals that does not exceed a given level of the IT security index. For example, at a glance, one will see that

62.17% (1071/1723) of hospitals have achieved an IT security index of 1.33 or less. In absolute terms, these levels seem very low. Is it possible that the less electronically advanced hospitals would deem it unnecessary to implement extended IT security and privacy practices, thus displaying lower levels of IT security index? To test this hypothesis, we confronted the IT security index with the transition level toward a fully electronic-based system.

Before performing this comparison, it is worthwhile to note that a majority of sampled hospitals (1056/1723; 61.28%) consider currently using a hybrid system (transition levels 4, 5, and 6), as they were more or less halfway toward a fully electronic-based system. An electronic-dominant system is found in 26.29% (453/1723) of hospitals, whereas a paper-dominant system is found in only 12.42% (214/1723) of them.

In Figure 1, we present the mean scores on our IT security index (vertical axis) for different groups of hospitals according to their level in transition toward a fully electronic-based system (horizontal axis).

Globally speaking, the mean score of IT security index is low for all 3 groups constituted on the basis of their transition level to a completely electronic system. It appears that the IT security index improves as one moves from the group of hospitals with a paper-dominant system (levels 1, 2, and 3) to the group of hospitals with electronic-dominant system (levels 7, 8, and 9) through the group of hospitals in a hybrid position (levels 4, 5, and 6). However, a further analysis depicts a much more nuanced picture. For the purposes of this analysis, we produced a scatter plot (Figure 2) indicating every hospital’s coordinate (a and b) depicting its position with regard to the transition level of its system (a) and to its IT security index (b). To facilitate the reading of the coordinates, we used a “jitter” to make sure that multiple observations that fall in the same coordinate can be visualized.

Table 3. Information technology security index level statistics.

IT ^a security index level	n (%)	Cumulative %
0.00	225 (13.06)	13.06
0.33	186 (10.80)	23.86
0.67	104 (6.04)	29.90
1.00	269 (15.61)	45.51
1.33	287 (16.66)	62.17
1.67	276 (16.02)	78.19
2.00	95 (5.51)	83.70
2.33	101 (5.86)	89.56
2.67	130 (7.54)	97.10
3.00	50 (2.90)	100.00
Total	1723 (100.00)	__ ^b

^aIT: information technology.

^bNot applicable.

Figure 1. Transition level toward electronic-based system versus information technology security index-1.

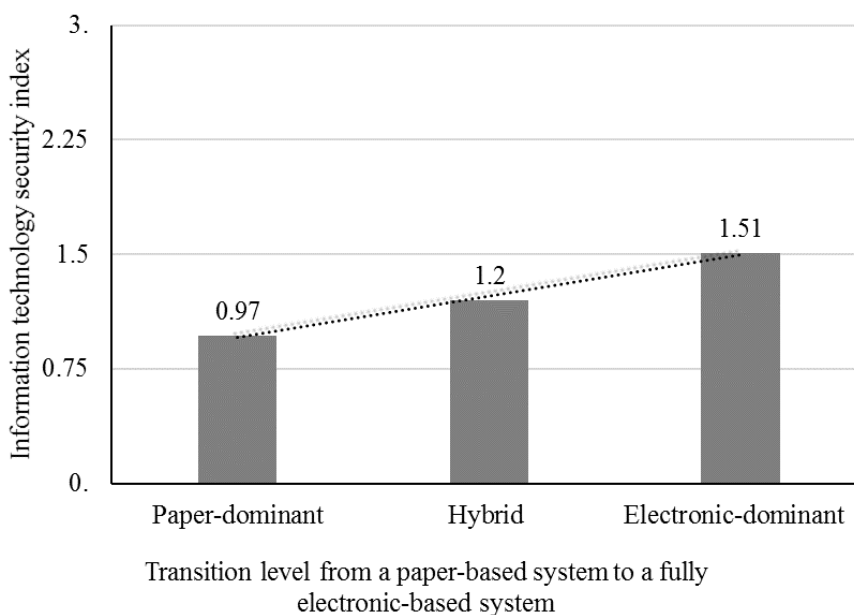
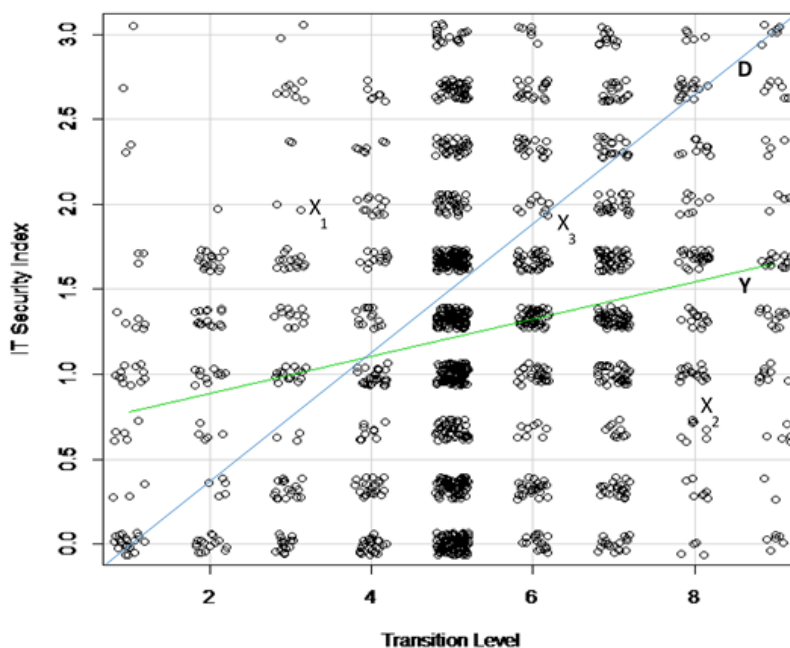


Figure 2. Transition level toward electronic-based system versus information technology security index-2. IT: information technology.



For facilitating the analysis of the figure, we added a diagonal line (D). If hospitals were enhancing their IT security and privacy practices as they moved forward, points representing hospitals would be scattered around the diagonal line. Instead, we found that points are scattered almost all over the surface of the figure.

Hospitals far above (far below) the diagonal line display a security index that is superior (inferior) to the average level required by their transition level toward a fully electronic-based system. For example, any hospital represented by X₁ point has an IT-related security index above the theoretical level required by its progress toward an electronic-dominant system. Conversely, a hospital represented by X₂ displays a security

index far below the level it should attain, considering how far it has progressed toward a fully electronic-based system. The security index of hospital X₃ is consistent with its progress in electronic-based system implementation.

The slope of the ascending curve in the figure (Y) suggests that there is a trend toward increasing IT security and privacy measures when hospitals move from a paper-based system to an electronic-based system. Though this finding is positive, it appears that the trend is not strong enough. Otherwise, the shape of the curve would be closer to the diagonal line.

Information Technology Security Level Versus Context Variables and Paper/Electronic Transition Level

In the previous section, we tested the relationship between the level of transition toward fully electronic systems on the one hand and the level of implementation of IT security on the other. We intuitively assumed a linear relationship between the 2. In this section, we take the analysis further, not only by considering the contextual variables in addition to the paper/electronic transition level but also by testing the validity of the linear relationship assumption between the transition level and IT security level.

In Table 4 we present the results of multivariate regression analyses testing whether hospitals' IT security levels can be predicted by hospitals' characteristics (contextual variables) and/or their transition level from paper to electronic system.

The results from both linear models (models 1 and 2) show that hospitals' characteristics and their transition level toward fully electronic-based system significantly contribute to the IT security level, with total R^2 equals, respectively, to 8.6% ($F_{16,1706}=10.05$; $P=.001$) and 12.0% ($F_{1,1705}=65.12$; $P=.001$). With models 3 and 4, we tested alternative nonlinear models and concluded that they do not bring about any significant contribution.

Table 4. Regression analyses (dependent variable: information technology security level).

Model	R	R^2	Adjusted R^2	Delta R^2	F (df)	P value
1 ^a	.294	.086	.078	.086	10.052 (16)	<.001
2 ^b	.346	.120	.111	.034	65.116 (1)	<.001
3 ^c	.346	.120	.111	.000	0.104 (1)	.75
4 ^d	.348	.121	.111	.001	1.911 (1)	.17

^aModel 1 predictors: contextual variables (see Table 1).

^bModel 2 predictors: contextual variables+ transition level toward electronic system.

^cModel 3 predictors: contextual variables+(transition level toward electronic system)².

^dModel 4 predictors: contextual variables+(transition level toward electronic system)³.

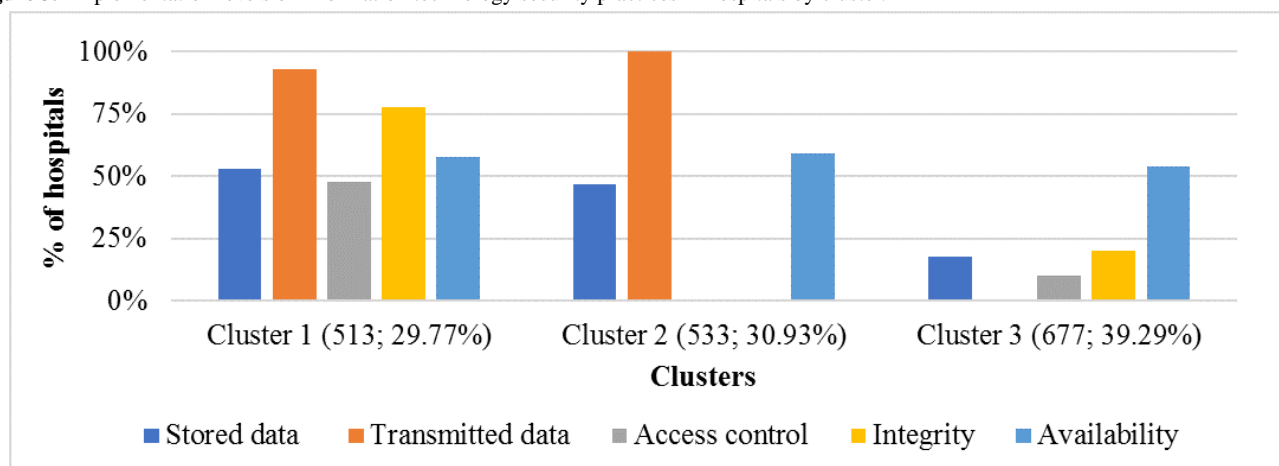
Table 5. Health information technology security patterns from cluster analysis. Within rows, different superscripts (*, †, and ‡) indicate significant ($P<.05$) pair-wise differences between means on Tamhane's T2 (post hoc) test.

Variable (grand mean)	Cluster			Analysis of variance	
	1, n (%)=513 (29.77)	2, n (%)=533 (30.93)	3, n (%)=677 (39.29)	F test (df)	P value
Confidentiality					
Stored data (0.37)	0.53*	0.47*	0.18†	106.6 (2)	<.001
Transmitted data (0.59)	0.93†	1.00*	0.00‡	9614.0 (2)	<.001
Access control (0.18)	0.48*	0.00‡	0.10†	303.9 (2)	<.001
Integrity (0.31)	0.78*	0.00‡	0.20†	763.7 (2)	<.001
Availability (0.57)	0.58	0.59	0.54	1.7 (2)	.18

Results of Cluster Analysis

The 3 patterns of HIT-related security practices resulting from our cluster analysis are presented in Table 5. The patterns in Table 5 are alternatively depicted in Figure 3, which shows cluster by cluster the implementation levels of each IT security practice measured in this study.

Before analyzing cluster differences, it is worth noting the grand mean of HIT security practices in sampled hospitals. As our security variables are measured through a dichotomous scale (1 if a practice is implemented and 0 if not implemented), the grand mean corresponds to the rate of hospitals that have a given practice implemented. This rate is presented in brackets in the column "variable" of Table 5. Overall, the most implemented practice is the one intended to ensure the confidentiality of electronically-transmitted data (present in 59% of hospitals), closely followed by the practice aiming at guaranteeing the availability of health data in case of a disaster (57%). The less implemented security practice is the access control or the IT workstations that contain sensitive health information (18%). This means that many hospitals tend to overlook the insider threat, which is preoccupying considering that insider threats represent over 50% of IT security breaches and are potentially more devastating than external threats [52,53].

Figure 3. Implementation levels of information technology security practices in hospitals by cluster.

On the basis of Tamhane's post hoc test (Table 5), we can immediately see that the "availability" criterion does not allow for the discrimination among the 3 clusters: hospitals that have implemented security measures allowing them to immediately recover their EHRs after a disaster are found in almost the same proportions in the 3 clusters (54% to 59%). However, there are differences among the 3 clusters with regard to confidentiality and integrity-related practices.

Overall, the strongest implementers of IT security practices are found in cluster 1, which accounts for 29.77% (513/1723) of surveyed hospitals. The implementation rate of each of the 4 distinctive HIT-related security practices is higher in this cluster than the average rate for all hospitals in our sample.

The weakest implementers of IT security practices are grouped in cluster 3, which comprises 39.29% (677/1723) of the sample. Implementation rates for all IT security practices are lower in this cluster than the overall average rates of the sample. Of note is that none of the 677 hospitals in this group uses encryption to protect electronically-transmitted health records, and only 10.0% of them enforce an access control to HIT systems.

In the middle position comes cluster 2 (533/1723; 30.93%). Although hospitals in cluster 2 have implemented IT security measures related to stored data and transmitted data in higher proportions, none of them have implemented either access control measures or integrity-related practices.

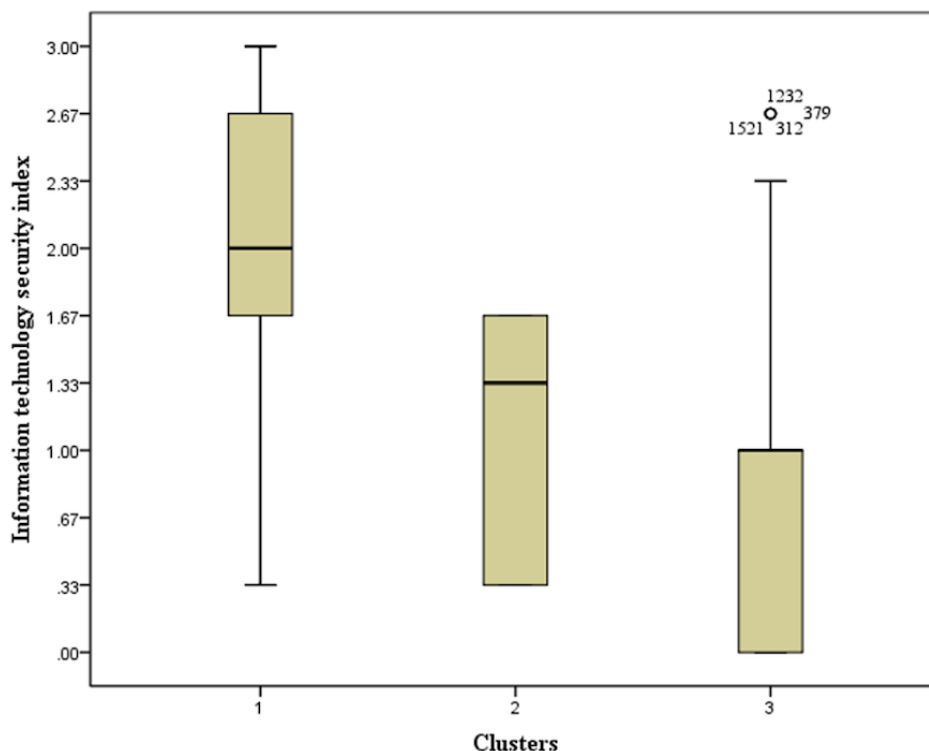
Discussion

Health Information Technology Security Index Across Clusters

The IT security index for all the hospitals in our sample (mean of 1.26) depicts a low level of the implementation of HIT

security practices, considering HIT users' concerns over IT-related security and privacy [7]. The IT security index level varies across the 3 clusters. It is relatively strong (2.00) for cluster 1, very low (0.83) for cluster 3, and low (1.08) for cluster 2. The box plot presented in Figure 4 shows the distribution of observations within each cluster according to the level of the IT security index. Moreover, 50% of hospitals in the strongest cluster (1) display an IT security index equal to or greater than 2, whereas for 25% of them, the IT security index is equal to or above 2.67. Cluster 3 is the only cluster that comprises hospitals with an IT security index that equals zero (25% of hospitals in this cluster). Furthermore, 75% of hospitals in cluster 3 present an IT security index that is equal to or less than 1. The IT security index for hospitals in cluster 2 varies from a minimum of 0.33 to a maximum of 1.67, with 50% of hospitals in this cluster displaying an IT security index between 1.33 and 1.67.

The results from the boxplot are meaningful in that they support our cluster qualification about each cluster's relative strength or weakness of HIT security implementation. It is particularly interesting to note that despite the gradation that establishes a hierarchy among clusters, allowing a stronger cluster to be identified compared with others in terms of IT security practices implemented, the overall picture does not look very good. The "strong" position of hospitals in cluster 1 is relative. In other words, the strongest hospitals in terms of IT security practices implementation appear so simply because others are badly failing.

Figure 4. Health information technology security index across clusters. IT: information technology.

Influence of Context Variables on Health Information Technology Security Patterns

To understand what determines HIT security patterns, we analyzed the contextual variables cluster by cluster. The aim here was to account for the influence of variables “theoretically related to the clusters, but not used in defining clusters” [49].

[Multimedia Appendix 2](#) presents the results of the test of independence (chi-square test of goodness of fit) carried out to assess whether characteristics of hospitals significantly vary according to cluster membership. We do so by comparing the observed and expected distributions of hospitals in different clusters, and the chi-square tests indicate whether or not the observed distributions significantly depart from the expected distribution.

From [Multimedia Appendix 2](#), one can note that only 2 out of 6 hospital characteristics (namely the university affiliation and the size of hospitals) do not allow in any way to significantly discriminate among hospitals’ membership in 1 cluster or another.

The remaining categories of contextual variables present at least 1 characteristic that is significantly overrepresented in 1 or 2 clusters and underrepresented elsewhere. Hospitals in the strongest cluster with regard to HIT security practices (cluster 1) are less likely to fall in the category of “independent/one site” hospitals, and they are underrepresented among hospitals that devote less than 1% of their total budget to the IT function. Conversely, hospitals in cluster 1 are more likely to fall into the category of “independent/multiple sites,” and they are more likely to rely on both national and regional security regulations.

On the other end of the security implementation spectrum, hospitals in the weakest cluster (cluster 3) are overrepresented in the category of “independent/one site” hospitals, as well as among hospitals that devote less than 1% of their total budget to the IT function. They are less likely to base their IT security and privacy practices on security regulations at any level (national, regional, and hospital).

The results in relation to the context factors provide important insights and raise some questions. The nonsignificance of the university affiliation status of hospitals goes against the expected results. Previous studies have shown that hospitals’ university affiliation is not neutral in their relationship to IT adoption [54] and health data breach risks [55]. This last study shows that health data breaches were more likely to happen in university-affiliated hospitals than in nonaffiliated hospitals. However, our results clearly indicate that the hospitals in our sample are spread across the different clusters (corresponding to different levels of IT security practices) regardless of their university affiliation status. Being at increased risk for health data breaches [55], university-affiliated hospitals should implement IT security measures that are commensurate with their risks.

The nonsignificance of the size of hospitals (number of beds) was also surprising, as one would expect large hospitals to be more aware of HIT security than smaller hospitals: we expected large hospitals to be significantly more represented in cluster 1, which is the strongest with regard to the implementation of HIT security practices. In addition, compared with small and medium-sized hospitals, large hospitals are associated with an increased risk of health data breaches [55], which is one more reason to strengthen their IT security practices.

Table 6. Transition toward an electronic-based system and health information technology security patterns.

Transition level	Number of hospitals									Chi-square (<i>df</i>)	<i>P</i> value
	Cluster 1, n (%)=513 (29.77)			Cluster 2, n (%)=533 (30.93)			Cluster 3, n (%)=677 (39.29)				
	O ^a	E ^b	R ^c	O	E	R	O	E	R		
Paper-dominant, n (%)=214 (12.42)	45	63.7	-18.7	60	66.2	-6.2	109	84.1	24.9	13.4	.001
Hybrid, n (%)=1056 (61.28)	302	314.4	-12.4	329	326.7	2.3	425	414.9	10.1	0.8	.69
Electronic-dominant, n (%)=453 (26.29)	166	134.9	31.1	144	140.1	3.9	143	178	-35.0	14.2	.001

^aO: observed.

^bE: expected.

^cR: residual.

The remaining contextual factors depict a more or less expected picture. The independence status (as opposed to being part of a group of hospitals/care institutions) and the number of sites on which a hospital operates seem to be good predictors of belonging to a weak or strong cluster in terms of implemented IT security practices. Hospitals operating on multiple sites or belonging to a group of other health care institutions need to apply tight IT security measures as their level of exposure is increased, and our results show that there is a trend in that direction. These factors (as well as the university affiliation and the hospital's size) are structural in nature, and health authorities cannot always act on them to affect IT security practices. However, by understanding their influence on the implementation of IT security practices, health authorities could consider more targeted measures.

Health authorities can act on the level of the hospitals' IT budget (as a percentage of total hospital budget). Our results suggest that a very low IT budget (in our case, an IT budget less than 1%) is not good for IT security practices implementation. Interestingly, they also suggest that a higher IT budget (more than 3.1%) does not make any difference. The optimal level of IT budget seems to be a level between 1% and 3%. This result raises more questions than answers. First, can we infer from this result that the relationship between the levels of IT budget and the levels of IT security practices implemented is in the shape of an inverted u-curve, meaning that very lower levels of IT budget are as counterproductive as very higher levels? Second, what factors would explain why the range between 1% and 3% is related to stronger IT practices implementation? This question is more challenging given that we do not know the portion of the overall IT budget hospitals in our sample specifically devote to IT security. These questions deserve more attention from researchers, and the answers would provide hospitals with guidance to avoid both under or overinvestment in IT security.

The security regulation constitutes another contextual factor that health authorities can use as a leverage to encourage the implementation of IT security practices in hospitals. It is interesting to note that the adoption of any of the 3 types of security regulations (national, regional, and hospital level) appears to be better (in terms of IT security practices

implemented) than nonadoption. However, national and regional regulations seem better than hospital-based security regulations.

Transition Toward an Electronic-Based System and Health Information Technology Security Patterns

Table 6 presents how hospitals at different phases in their transition toward a fully electronic-based system are distributed across the 3 clusters depicting HIT security patterns. It appears that hospitals in the earlier phases of the transition (paper-dominant hospitals) are less likely to belong to cluster 1 and more likely to belong to cluster 3. Conversely, hospitals that are well-advanced toward electronic-based systems (electronic-dominant hospitals) are more likely to belong to cluster 1 and less likely to belong to cluster 3. On the surface, this seems good news as digitized hospitals are more likely to be found in the strongest cluster with regards to HIT security and privacy practices. However, a closer look at the distribution of hospitals that have an electronic-dominant system across the 3 clusters leads to a less optimistic conclusion: actually, only 36.6% (166 out of 453) of these hospitals belong to cluster 1. This means that there are many hospitals that consider themselves to be well-advanced in their digitization process and that at the same time display major weaknesses in their IT security and privacy practices.

These results are a cause of concern when one considers (1) that the top-ranking risk associated with EHR is the "privacy of data - access control [56]" and (2) that some patients are reluctant to disclose their health information to protect themselves against the perceived EHR privacy and security risks [57].

Ultimately, the consequences of the relative weaknesses in IT security and privacy practices will obscure the benefits expected from the European hospitals' transition toward fully electronic-based systems.

Implications and Conclusion

This study highlights a disturbing state in European hospitals regarding the level of implementation of HIT security and privacy practices. Overall, none of the 5 basic security practices investigated in this study is present in more than 60% of surveyed hospitals. Moreover, 3 out of 5 practices are absent

in more than two-thirds of the hospitals that were surveyed in this study. These statistics are unsettling as security practices studied here are basic practices that should be implemented in almost all hospitals. Encryption for stored data is used in only 37% of hospitals. It is used for transmitted data in only 59% of hospitals. Many hospitals (more than 80%) do not deem it necessary to control the access to workstations containing health data with health professionals' cards or codes. There is as few as 18% of hospitals that have implemented these practices. Hospitals in which all these measures are not implemented expose health information to a breach of confidentiality.

In this study, the practices related to integrity and availability are respectively measured at 31% and 57% implementation rates in hospitals. These implementation rates are low for systems containing highly sensitive information. They mean that (1) in almost 70% hospitals, health data in IT systems can be modified by nonauthorized persons, provided they have access to the systems, and (2) more than 40% of surveyed European hospitals would not be able to restore critical clinical information in the aftermath of an incident, resulting in a partial or complete loss of data.

There is another way of looking at our results. Our cluster analysis allowed us to identify 3 patterns of HIT security practices. The majority of surveyed hospitals falls into the 2 worst clusters (clusters 2 and 3): these 2 clusters total 1210 hospitals out of 1723 (70.22%). This means that 7 out of 10 European hospitals are performing poorly in ensuring the security of their EHR.

We expected that hospitals that are well-advanced in their transition toward a fully electronic health system would display higher levels of implementation of IT security practices. Confronting each hospital's security index (a compounded measure of implemented IT security and privacy practices) to its level of transition toward a fully electronic health system, we unraveled a rather mixed situation, showing that there are many electronically advanced hospitals that are poor implementers of IT security practices. This is great concern not only for hospitals that are poorly equipped in IT security but also for other hospitals with which they share health data and above all for patients whose health information privacy is not adequately protected.

The results presented in this study have theoretical and practical implications. From a theoretical standpoint, it would be helpful to further investigate why the health care sector continues to lag behind in terms of IT security practices implementation. How can one explain the contradiction between the stated importance of IT security measures in hospital settings and the weaknesses in this regard? Are there any sector-related factors that explain the poor implementation of IT security practices

in health care organizations? Another research avenue stems from 1 of the limitations of this study. As already mentioned in the Methods section, in the absence of sound evidence suggesting unequal importance of the 3 components of the IT security triad (confidentiality, integrity, and availability), we have assumed an equal weight for each of them. Future research could challenge this assumption and provide empirical-based evidence of the relative importance of the 3 components of the IT security measure. Future research could also determine whether and to what extent this relative importance depends on organizational context. Our results about some context variables, namely the university affiliation status and the size of hospitals, raised an interesting question: why hospitals that present a higher risk profile (university-affiliated hospitals and large hospitals), according to previous research [55], do not appear to be among the hospitals that take relatively strong IT security measures? This question represents an interesting research avenue. Another interesting research avenue is about the determination of the optimal level of IT security investment in hospital settings.

From a practical standpoint, this study raises a red flag that hospitals managers and health policy makers should not ignore. The replacement of paper-based systems with electronic-based systems comes with increased IT-related security risks and requires adequate IT security measures. Thus, hospital managers should make sure to include IT-security practices in their plans toward fully electronic-based systems. For health policy makers, when designing incentives for meaningful use of IT in health care organizations, they should include stringent IT security requirements. Both hospital managers and health policy makers should monitor the digitization process of hospitals to ensure that the implementation of IT security practices keeps pace with the increasing usage of IT.

Although we had access to an interesting dataset collected through a survey by the European Union, we were limited to the questions asked in the survey. This is the problem of using secondary data. We also acknowledge some limits stemming from our definition of security practices. One could enlarge this definition or completely choose other security practices. Besides, we made the assumption of equal weight for all the 3 components (confidentiality, integrity, and availability) of the IT security measure. This is a limitation the reader might take into account when interpreting our results. For the transition level toward a fully electronic-based system, we relied on a self-reported level given by each hospital's IT manager in the absence of a more objective measure. This can be somehow biased. Despite the abovementioned limits, this paper contributes to the understanding of IT security practices in health care organizations. It also contributes to raising awareness of the security issues that can impede the effective delivery of health care services.

Authors' Contributions

The 2 first authors contributed to the conception and design of the study. All authors contributed to the data analysis, interpretation, and discussion of the results and reviewed and approved this version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Correlations matrix.

[\[PDF File \(Adobe PDF File\), 65KB - medinform_v7i1e11211_app1.pdf\]](#)

Multimedia Appendix 2

Influence of context variables on health information technology security patterns.

[\[PDF File \(Adobe PDF File\), 160KB - medinform_v7i1e11211_app2.pdf\]](#)

References

1. Agrawal R, Grandison T, Johnson C, Kiernan J. Enabling the 21st century health care information technology revolution. *Commun ACM* 2007 Feb 1;50(2):34-42. [doi: [10.1145/1216016.1216018](#)]
2. Bitton A, Flier LA, Jha AK. Health information technology in the era of care delivery reform: to what end? *J Am Med Assoc* 2012 Jun 27;307(24):2593-2594. [doi: [10.1001/jama.2012.6663](#)] [Medline: [22735426](#)]
3. Sheikh A, Sood HS, Bates DW. Leveraging health information technology to achieve the "triple aim" of healthcare reform. *J Am Med Inform Assoc* 2015 Jul;22(4):849-856 [FREE Full text] [doi: [10.1093/jamia/ocv022](#)] [Medline: [25882032](#)]
4. Windle JR, Windle TA. Electronic health records and the quest to achieve the "triple aim". *J Am Coll Cardiol* 2015 May 12;65(18):1973-1975 [FREE Full text] [doi: [10.1016/j.jacc.2015.03.038](#)] [Medline: [25953749](#)]
5. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. *J Med Internet Res* 2018 Feb 5;20(2):e41 [FREE Full text] [doi: [10.2196/jmir.8793](#)] [Medline: [29402759](#)]
6. Tejero A, de la Torre I. Advances and current state of the security and privacy in electronic health records: survey from a social perspective. *J Med Syst* 2012 Oct;36(5):3019-3027. [doi: [10.1007/s10916-011-9779-x](#)] [Medline: [21947856](#)]
7. Shaw N, Kulkarni A, Mador RL. Patients and health care providers' concerns about the privacy of electronic health records: a review of the literature. *Electron J Health Informatics* 2011;6(1) [FREE Full text]
8. Miller AR, Tucker C. Privacy protection and technology diffusion: The case of electronic medical records. *Manag Sci* 2009 Jul;55(7):1077-1093. [doi: [10.1287/mnsc.1090.1014](#)]
9. Jones SS, Rudin R, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. *Ann Intern Med* 2014 Jan 7;160(1):48-54. [doi: [10.7326/M13-1531](#)] [Medline: [24573664](#)]
10. Bahtiyar S, Çağlayan MU. Trust assessment of security for e-health systems. *Electron Commer Res Appl* 2014;13(3):164-177. [doi: [10.1016/j.elerap.2013.10.003](#)]
11. Kisekka V, Giboney J. The effectiveness of health care information technologies: evaluation of trust, security beliefs, and privacy as determinants of health care outcomes. *J Med Internet Res* 2018 Apr 11;20(4):e107 [FREE Full text] [doi: [10.2196/jmir.9014](#)] [Medline: [29643052](#)]
12. Vogel J, Brown JS, Land T, Platt R, Klompas M. MDPHnet: secure, distributed sharing of electronic health record data for public health surveillance, evaluation, and planning. *Am J Public Health* 2014 Dec;104(12):2265-2270. [doi: [10.2105/AJPH.2014.302103](#)] [Medline: [25322301](#)]
13. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. *J Med Internet Res* 2018 May 29;20(5):e185 [FREE Full text] [doi: [10.2196/jmir.9134](#)] [Medline: [29844010](#)]
14. Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. *J Med Internet Res* 2016 Apr 15;18(4):e66 [FREE Full text] [doi: [10.2196/jmir.5011](#)] [Medline: [27083521](#)]
15. Adler-Milstein J, Ronchi E, Cohen GR, Winn LA, Jha AK. Benchmarking health IT among OECD countries: better data for better policy. *J Am Med Inform Assoc* 2014;21(1):111-116 [FREE Full text] [doi: [10.1136/amiajnl-2013-001710](#)] [Medline: [23721983](#)]
16. Häyrynen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *Int J Med Inform* 2008 May;77(5):291-304. [doi: [10.1016/j.ijmedinf.2007.09.001](#)] [Medline: [17951106](#)]
17. Achampong E. Electronic health record (EHR) and cloud security: the current issues. *IJ-CLOSER* 2014 Feb 13;2(6):417-420. [doi: [10.11591/closer.v2i6.5343](#)]
18. Rodrigues J, de la Torre I, Fernández G, López-Coronado M. Analysis of the security and privacy requirements of cloud-based electronic health records systems. *J Med Internet Res* 2013 Aug 21;15(8):e186 [FREE Full text] [doi: [10.2196/jmir.2494](#)] [Medline: [23965254](#)]
19. Burns AJ, Johnson ME. Securing health information. *IT Prof* 2015 Jan;17(1):23-29. [doi: [10.1109/MITP.2015.13](#)]

20. Koehler N, Vujovic O, McMenamin C. Healthcare professionals' use of mobile phones and the internet in clinical practice. *J Mob Technol Med* 2013;2(1):3-13. [doi: [10.7309/jmtm.2.1.2](https://doi.org/10.7309/jmtm.2.1.2)]
21. Williams J. Left to their own devices how healthcare organizations are tackling the BYOD trend. *Biomed Instrum Technol* 2014;48(5):327-339. [doi: [10.2345/0899-8205-48.5.327](https://doi.org/10.2345/0899-8205-48.5.327)] [Medline: [25244198](https://pubmed.ncbi.nlm.nih.gov/25244198/)]
22. Dehling T, Gao F, Schneider S, Sunyaev A. Exploring the far side of mobile health: information security and privacy of mobile health apps on iOS and Android. *JMIR Mhealth Uhealth* 2015 Jan 19;3(1):e8-17 [FREE Full text] [doi: [10.2196/mhealth.3672](https://doi.org/10.2196/mhealth.3672)] [Medline: [25599627](https://pubmed.ncbi.nlm.nih.gov/25599627/)]
23. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med* 2015 Sep 7;13:214 [FREE Full text] [doi: [10.1186/s12916-015-0444-y](https://doi.org/10.1186/s12916-015-0444-y)] [Medline: [26404673](https://pubmed.ncbi.nlm.nih.gov/26404673/)]
24. Dimitropoulos L, Patel V, Scheffler SA, Posnack S. Public attitudes toward health information exchange: perceived benefits and concerns. *Am J Manag Care* 2011 Dec;17(12 Spec No):SP111-SP116 [FREE Full text] [Medline: [22216769](https://pubmed.ncbi.nlm.nih.gov/22216769/)]
25. Zhang A, Bacchus A, Lin X. Consent-based access control for secure and privacy-preserving health information exchange. *Security Comm Networks* 2016 Jul 7;9(16):3496-3508 [FREE Full text] [doi: [10.1002/sec.1556](https://doi.org/10.1002/sec.1556)]
26. Paganini P. Infosec Institute. 2014. Risks and cyber threats to the healthcare industry URL: <https://resources.infosecinstitute.com/risks-cyber-threats-healthcare-industry/> [accessed 2018-06-01] [WebCite Cache ID 6zrHNa93B]
27. Healthcare Information Security. Princeton, NJ: ISMG; 2014. The State of Healthcare Information Security Today. Update on HIPAA Omnibus Compliance, Protecting Patient Data URL: <https://www.healthcareinfosecurity.com/surveys/state-healthcare-information-security-today-s-23> [accessed 2019-02-04] [WebCite Cache ID 75veAC2ft]
28. HIMSS. Chicago, IL: HIMSS; 2015 Jun. 2015 HIMSS Cybersecurity Survey URL: <https://www.himss.org/2015-cybersecurity-survey/full-report> [accessed 2019-02-04] [WebCite Cache ID 75veOqmSy]
29. Ma Q, Schmidt MB, Pearson JM, Herberger GR. An integrated framework for information security management. *Rev Bus* 2009;30(1):58-69.
30. Absolute Software Corporation. Austin, Texas; 2015. The Cost of a Data Breach: Healthcare Settlements Involving Lost or Stolen Devices URL: <https://www.absolute.com/en/resources/whitepapers/cost-of-a-healthcare-data-breach> [accessed 2019-02-04] [WebCite Cache ID 75vet1URB]
31. von Solms SH. Information security governance: compliance management vs operational management. *J Comput Security* 2005 Sep;24(6):443-447. [doi: [10.1016/j.cose.2005.07.003](https://doi.org/10.1016/j.cose.2005.07.003)]
32. Pfleeger CP, Pfleeger SL, Margulies J. Security in computing. In: *Security In Computing (5th Edition)*. Upper Saddle River, NJ: Prentice Hall; Feb 5, 2015:944.
33. Dehling T, Sunyaev A. Secure provision of patient-centered health information technology services in public networks—leveraging security and privacy features provided by the German nationwide health information technology infrastructure. *Electron Markets* 2014 Feb 8;24(2):89-99 [FREE Full text] [doi: [10.1007/s12525-013-0150-6](https://doi.org/10.1007/s12525-013-0150-6)] [Medline: [23881944](https://pubmed.ncbi.nlm.nih.gov/23881944/)]
34. Brumen B, Heričko M, Sevčnikar A, Završnik J, Hölbl M. Outsourcing medical data analyses: can technology overcome legal, privacy, and confidentiality issues? *J Med Internet Res* 2013 Dec 16;15(12):e283 [FREE Full text] [doi: [10.2196/jmir.2471](https://doi.org/10.2196/jmir.2471)] [Medline: [24342053](https://pubmed.ncbi.nlm.nih.gov/24342053/)]
35. European Commission. Luxembourg: Publications Office of the European Union; 2014 Mar 24. European Hospital Survey: Benchmarking Deployment of eHealth Services (2012-2013): JRC Scientific and Policy Reports URL: <https://ec.europa.eu/digital-single-market/en/news/european-hospital-survey-benchmarking-deployment-ehealth-services-2012-2013> [accessed 2019-02-04] [WebCite Cache ID 75vfRkPjv]
36. Kwon J, Johnson ME. Proactive versus reactive security investments in the healthcare sector. *MIS Q* 2014 Feb 2;38(2):451-471. [doi: [10.25300/MISQ/2014/38.2.06](https://doi.org/10.25300/MISQ/2014/38.2.06)]
37. Coventry L, Branley D. Cybersecurity in healthcare: a narrative review of trends, threats and ways forward. *Maturitas* 2018 Jul;113:48-52. [doi: [10.1016/j.maturitas.2018.04.008](https://doi.org/10.1016/j.maturitas.2018.04.008)] [Medline: [29903648](https://pubmed.ncbi.nlm.nih.gov/29903648/)]
38. McLeod A, Dolezel D. Cyber-analytics: modeling factors associated with healthcare data breaches. *Decis Support Syst* 2018 Apr;108:57-68. [doi: [10.1016/j.dss.2018.02.007](https://doi.org/10.1016/j.dss.2018.02.007)]
39. Kruse CS, Frederick B, Jacobson T, Monticone DK. Cybersecurity in healthcare: a systematic review of modern threats and trends. *Technol Health Care* 2017;25(1):1-10. [doi: [10.3233/THC-161263](https://doi.org/10.3233/THC-161263)] [Medline: [27689562](https://pubmed.ncbi.nlm.nih.gov/27689562/)]
40. Lopes IM, Oliveira P. Implementation of the general data protection regulation: A survey in health clinics. 2018 Presented at: 13th Iberian Conference on Information Systems and Technologies (CISTI); June 13-16, 2018; Cáceres, Spain p. 1-6. [doi: [10.23919/CISTI.2018.8399156](https://doi.org/10.23919/CISTI.2018.8399156)]
41. Angst CM, Block ES, D'Arcy J, Kelley K. When do IT security investments matter? Accounting for the influence of institutional factors in the context of healthcare data breaches. *MIS Q* 2017 Mar 3;41(3):893-916. [doi: [10.25300/MISQ/2017/41.3.10](https://doi.org/10.25300/MISQ/2017/41.3.10)]
42. Begun JW, Thygeson M. Managing complex healthcare organizations. In: Fottler MD, Malvey D, Slovensky DJ, editors. *Handbook of healthcare management*. Northampton: Edward Elgar; 2015:1-17.

43. Khan S, Vandermorris A, Shepherd J, Begun JW, Lanham HJ, Uhl-Bien M, et al. Embracing uncertainty, managing complexity: applying complexity thinking principles to transformation efforts in healthcare systems. *BMC Health Serv Res* 2018 Mar 21;18(1):192 [FREE Full text] [doi: [10.1186/s12913-018-2994-0](https://doi.org/10.1186/s12913-018-2994-0)] [Medline: [29562898](https://pubmed.ncbi.nlm.nih.gov/29562898/)]
44. Deidda M, Lupiáñez-Villanueva F, Maghiros I. European Commission - Joint Research Centre. Luxembourg; 2013. European hospital survey: Benchmarking deployment of e-health services (2012-2013), Methodological report URL: <https://ec.europa.eu/digital-single-market/en/news/european-hospital-survey-benchmarking-deployment-ehealth-services-2012-2013> [accessed 2019-02-04] [WebCite Cache ID 75vgmJUeu]
45. European Commission - Joint Research Centre. Luxembourg: European Commission - Joint Research Centre; 2014. European hospital survey: Benchmarking deployment of e-health services (2012-2013), Final report URL: <https://ec.europa.eu/digital-single-market/en/news/european-hospital-survey-benchmarking-deployment-ehealth-services-2012-2013> [accessed 2019-02-04] [WebCite Cache ID 75vhBRBsv]
46. Balijepally V, Mangalaraj G, Iyengar K. Are we wielding this hammer correctly? A reflective review of the application of cluster analysis in information systems research. *J Assoc Inf Syst* 2011 May;12(5):375-413. [doi: [10.17705/1jais.00266](https://doi.org/10.17705/1jais.00266)]
47. Hair JF, Babin B, Anderson RE, Black B, Tatham RL. *Multivariate Data Analysis 6th Edition*. Upper Saddle River, NJ: Prentice Hall; Oct 28, 2005.
48. Jung Y, Park H, Du DZ, Drake B. A decision criterion for the optimal number of clusters in hierarchical clustering. *J Global Optim* 2003 Jan;25(1):91-111. [doi: [10.1023/A:1021394316112](https://doi.org/10.1023/A:1021394316112)]
49. Ketchen DJ, Shook CL. The application of cluster analysis in strategic management research: an analysis and critique. *Strat Mgmt J* 1996 Jun;17(6):441-458. [doi: [10.1002/\(SICI\)1097-0266\(199606\)17:6<441::AID-SMJ819>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G)]
50. Kwon J, Johnson ME. Security practices and regulatory compliance in the healthcare industry. *J Am Med Inform Assoc* 2013 Jan 1;20(1):44-51 [FREE Full text] [doi: [10.1136/amiajnl-2012-000906](https://doi.org/10.1136/amiajnl-2012-000906)] [Medline: [22955497](https://pubmed.ncbi.nlm.nih.gov/22955497/)]
51. Hair JF, Black WC, Babin BJ. *Multivariate Data Analysis: A Global Perspective 7th Edition*. Upper Saddle River, New Jersey: Prentice Hall; 2010:800.
52. Spears JL, Barki H. User participation in information systems security risk management. *MIS Q* 2010 Sep;34(3):503-522. [doi: [10.2307/25750689](https://doi.org/10.2307/25750689)]
53. Safa NS, Maple C, Watson T, von Solms R. Motivation and opportunity based model to reduce information security insider threats in organisations. *J Inf Secur Appl* 2018 Jun;40:247-257. [doi: [10.1016/j.jisa.2017.11.001](https://doi.org/10.1016/j.jisa.2017.11.001)]
54. Poba-Nzaou P, Uwizeyemungu S, Raymond L, Paré G. Motivations underlying the adoption of ERP systems in healthcare organizations: insights from online stories. *Inf Syst Front* 2012 Jul 13;16(4):591-605. [doi: [10.1007/s10796-012-9361-1](https://doi.org/10.1007/s10796-012-9361-1)]
55. Gabriel MH, Noblin A, Rutherford A, Walden A, Cortelyou-Ward K. Data breach locations, types, and associated characteristics among US hospitals. *Am J Manag Care* 2018 Feb;24(2):78-84 [FREE Full text] [Medline: [29461854](https://pubmed.ncbi.nlm.nih.gov/29461854/)]
56. Thakkar M, Davis D. Risks, barriers, and benefits of EHR systems: a comparative study based on size of hospital. *Perspect Health Inf Manag* 2006 Aug 14;3:5 [FREE Full text] [Medline: [18066363](https://pubmed.ncbi.nlm.nih.gov/18066363/)]
57. Campos-Castillo C, Anthony DL. The double-edged sword of electronic health records: implications for patient disclosure. *J Am Med Inform Assoc* 2015 Apr;22(e1):e130-e140. [doi: [10.1136/amiajnl-2014-002804](https://doi.org/10.1136/amiajnl-2014-002804)] [Medline: [25059953](https://pubmed.ncbi.nlm.nih.gov/25059953/)]

Abbreviations

- C_{max}**: maximum chance criteria
C_{pro}: proportional chance criteria
eHealth: electronic health
EHR: electronic health record
HIE: health information exchange
HIT: health information technology
ISMG: Information Security Media Group
IT: information technology

Edited by G Eysenbach; submitted 02.06.18; peer-reviewed by M Abdelhamid, S Albakri, B Brumen; comments to author 06.10.18; revised version received 29.11.18; accepted 29.12.18; published 25.03.19.

Please cite as:

Uwizeyemungu S, Poba-Nzaou P, Cantinotti M
 European Hospitals' Transition Toward Fully Electronic-Based Systems: Do Information Technology Security and Privacy Practices Follow?
JMIR Med Inform 2019;7(1):e11211
 URL: <http://medinform.jmir.org/2019/1/e11211/>
 doi:[10.2196/11211](https://doi.org/10.2196/11211)
 PMID:[30907732](https://pubmed.ncbi.nlm.nih.gov/30907732/)

©Sylvestre Uwizeyemungu, Placide Poba-Nzaou, Michael Cantinotti. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Data Analysis and Visualization of Newspaper Articles on Thirdhand Smoke: A Topic Modeling Approach

Qian Liu^{1,2}, BTech, MSc, PhD; Qiuyi Chen¹, BA; Jiayi Shen³, MBBS; Huailiang Wu³, MBBS; Yimeng Sun¹, BA; Wai-Kit Ming^{4,5,6,7}, MD, MPH, MMSCI, PhD

¹School of Journalism and Communication, Jinan University, Guangzhou, China

²National Media Experimental Teaching Demonstration Center, Jinan University, Guangzhou, China

³School of Medicine, Jinan University, Guangzhou, China

⁴Department of Obstetrics & Gynaecology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

⁵Harvard Medical School, Harvard University, Boston, MA, United States

⁶Faculty of Medicine, The Chinese University of Hong Kong, Shatin, NT, China (Hong Kong)

⁷Li Ka Shing Faculty of Medicine, University of Hong Kong, Pokfulam, China (Hong Kong)

Corresponding Author:

Wai-Kit Ming, MD, MPH, MMSCI, PhD

Department of Obstetrics & Gynaecology

The First Affiliated Hospital, Sun Yat-sen University

58 Zhongshan 2nd Rd

Guangzhou,

China

Phone: 86 14715485116

Email: wkming@connect.hku.hk

Abstract

Background: Thirdhand smoke has been a growing topic for years in China. Thirdhand smoke (THS) consists of residual tobacco smoke pollutants that remain on surfaces and in dust. These pollutants are re-emitted as a gas or react with oxidants and other compounds in the environment to yield secondary pollutants.

Objective: Collecting media reports on THS from major media outlets and analyzing this subject using topic modeling can facilitate a better understanding of the role that the media plays in communicating this health issue to the public.

Methods: The data were retrieved from the Wisser and Factiva news databases. A preliminary investigation focused on articles dated between January 1, 2013, and December 31, 2017. Use of Latent Dirichlet Allocation yielded the top 10 topics about THS. The use of the modified LDAvis tool enabled an overall view of the topic model, which visualizes different topics as circles. Multidimensional scaling was used to represent the intertopic distances on a two-dimensional plane.

Results: We found 745 articles dated between January 1, 2013, and December 31, 2017. The United States ranked first in terms of publications (152 articles on THS from 2013-2017). We found 279 news reports about THS from the Chinese media over the same period and 363 news reports from the United States. Given our analysis of the percentage of news related to THS in China, Topic 1 (Cancer) was the most popular among the topics and was mentioned in 31.9% of all news stories. Topic 2 (Control of quitting smoking) was related to roughly 15% of news items on THS.

Conclusions: Data analysis and the visualization of news articles can generate useful information. Our study shows that topic modeling can offer insights into understanding news reports related to THS. This analysis of media trends indicated that related diseases, air and particulate matter (PM_{2.5}), and control and restrictions are the major concerns of the Chinese media reporting on THS. The Chinese press still needs to consider fuller reports on THS based on scientific evidence and with less focus on sensational headlines. We recommend that additional studies be conducted related to sentiment analysis of news data to verify and measure the influence of THS-related topics.

(*JMIR Med Inform* 2019;7(1):e12414) doi:[10.2196/12414](https://doi.org/10.2196/12414)

KEYWORDS

media concerns; topic modeling; third-hand smoke; tobacco; indoor air quality

Introduction

Thirdhand smoke (THS) is an important public health issue and has been an increasingly popular topic for decades in China since its first mention in 2009 [1]. Aggregating media reports from major media outlets and analyzing the coverage using topic modeling may help shed light on the role that the media plays in communicating this health concept. Thirdhand smoke consists of residual tobacco smoke pollutants that remain on surfaces and in dust after tobacco has been smoked. These pollutants are re-emitted into the gas phase (ie, off-gassing) or react with oxidants and other compounds in the environment to yield secondary pollutants [2]. Evidence supports the presence of THS in indoor environments. Thirdhand smoke is found in enclosed spaces where habitual smoking occurs, such as residences and automobiles [3]. This phenomenon is associated with health hazards: research has shown that residual nicotine from tobacco smoke absorbed onto indoor surfaces reacts with ambient nitrous acid to form carcinogenic tobacco-specific nitrosamines, which can cause significant levels of DNA damage in human cell lines [3]. The related acute and long-term risks of THS include disease and premature mortality. Children and infants are particularly susceptible to THS exposure [4].

Given the increasing interest in THS, the mass media has focused on delivering and communicating information on this topic to public audiences. Both in China and abroad, there have been media reports related to THS. According to search results of news articles mentioning THS, the United States ranks first in terms of topic mentions. However, previous studies have shown that fewer people are aware of the harms of THS than that of secondhand smoke [5]. Because the mass media is a key player in communicating health-related information, it could play a positive role in helping the public understand the risks of THS and the ways to protect themselves from it [6]. We therefore decided to compare this topic between China and the United States.

Multimodal data modeling combines information from different resources. Topic modeling refers to statistical models in which unstructured data are structured in accordance with latent themes to deal with multimodal data. Latent Dirichlet Allocation (LDA) is the most popular form of topic modeling and is a generative probabilistic modeling method for converting visual words into images and visual word documents [7-9]. Topic modeling has broad applications in various fields such as text mining [10], medicine [11-13], economics [14], and social network analysis [15]. To the best of our knowledge, however, there have been very few studies using LDA to evaluate the media's treatment of THS. As a result, we used LDA modeling method for our analysis despite its' being a common method. This paper aims to determine the current patterns and the role of mass communication related to THS.

Methods

Data Collection

We used the Wiser database for Chinese news content (from the Wise News website) and the Factiva database (from the Dow Jones website) to retrieve the international news articles.

The Wiser database is an ever-growing, professional Chinese media content database that contains more than one million data entries. It is currently the best source for Chinese media content given its large volume of data and timely updates. Factiva is an international news and information database that includes nearly 33,000 premium sources such as licensed publications, influential websites, blogs, images, and videos. To get a general idea of the topic of THS, we conducted a preliminary investigation of the Factiva database and retrieved articles dated between January 1, 2013, and December 31, 2017. The Wiser database includes data published only within the last five years. However, many of the articles we flagged merely noted the topic of smoking without further elaborating on the issue of THS. We next narrowed down the entries to only those newspaper sources pertaining to THS; we believe that articles from newspapers are much more reliable than those from other sources. We used LDA to further analyze the Chinese news articles.

An LDA topic model is a model with a three-level hierarchical Bayesian model. The basic assumption of this model is a combination of words belonging to different topics [16]. LDA suggests that there may be multiple topics in an article and that the wording in that article or paper reflects the exact set of topics that the reporter wished to address. Using Gibbs Sampling techniques, a method that estimates the marginal distribution of interested variables, we can determine the topics among the data pool [17].

Processing

Data processing (Figure 1) was conducted before applying LDA modeling by using Python to do data cleaning and the Python package, Jieba, to do the segmentation [18,19]. First, the redundant and null data were chosen to be removed, followed by removal of irrelevant information such as advertisements. Next, word segmentation was conducted using the Jieba package. However, THS articles with a lot of terminology were calculated and added based on the calculation of the words' information entropy and term frequency to avoid the influence of unprofessional dictionaries. New terms such as "China Anticancer Association," "Family Doctor," "Shenzhen Municipal Government," "Air Circulation," "Air Purification," "Chinese Preventive Medicine Association," "Tobacco Monopoly," "Tobacco Market," and so on, were added to the dictionary for further analysis. Furthermore, common Chinese stop words were removed such as "a," "of," "it," etc. Next, a document-term matrix was built, and term frequency-inverse document frequency (TF-IDF) was used in the data processing.

Multiple studies have been conducted related to the choice of LDA topic number and the explanations of each topic. In previous research related to topic number, 10-30 topics were assigned nearly the same log-likelihood measure. Therefore, we adopted the number 10 for the topic parameter. By analyzing key words, topic content was generated accordingly. If we considered only the statistical measures, the results might not be interpretable by humans [20]. Therefore, by combining both statistical measures and manual interpretation, we selected 10 topics to analyze by using Python version 3.6.1 with the LDAvis tool [16]. We set $\lambda=1$ and recovered 10 topics and their

keywords (Table 1). To elaborate on the topics, the topic name was generated based on the given keywords (Table 1) as well.

We sought topics that overlapped as in the visualization shown in Figures 2 and 3. In this two-dimensional plane, topics are

represented as cycles whose centers are determined by the computed topics' distances [16]. We categorized the topics into three main primary groups: links to related diseases, air and particulate matter (PM_{2.5}), and control and restrictions (Table 1).

Figure 1. Data processing chart. LDA: Latent Dirichlet Allocation; TF-IDF: frequency-inverse document frequency.

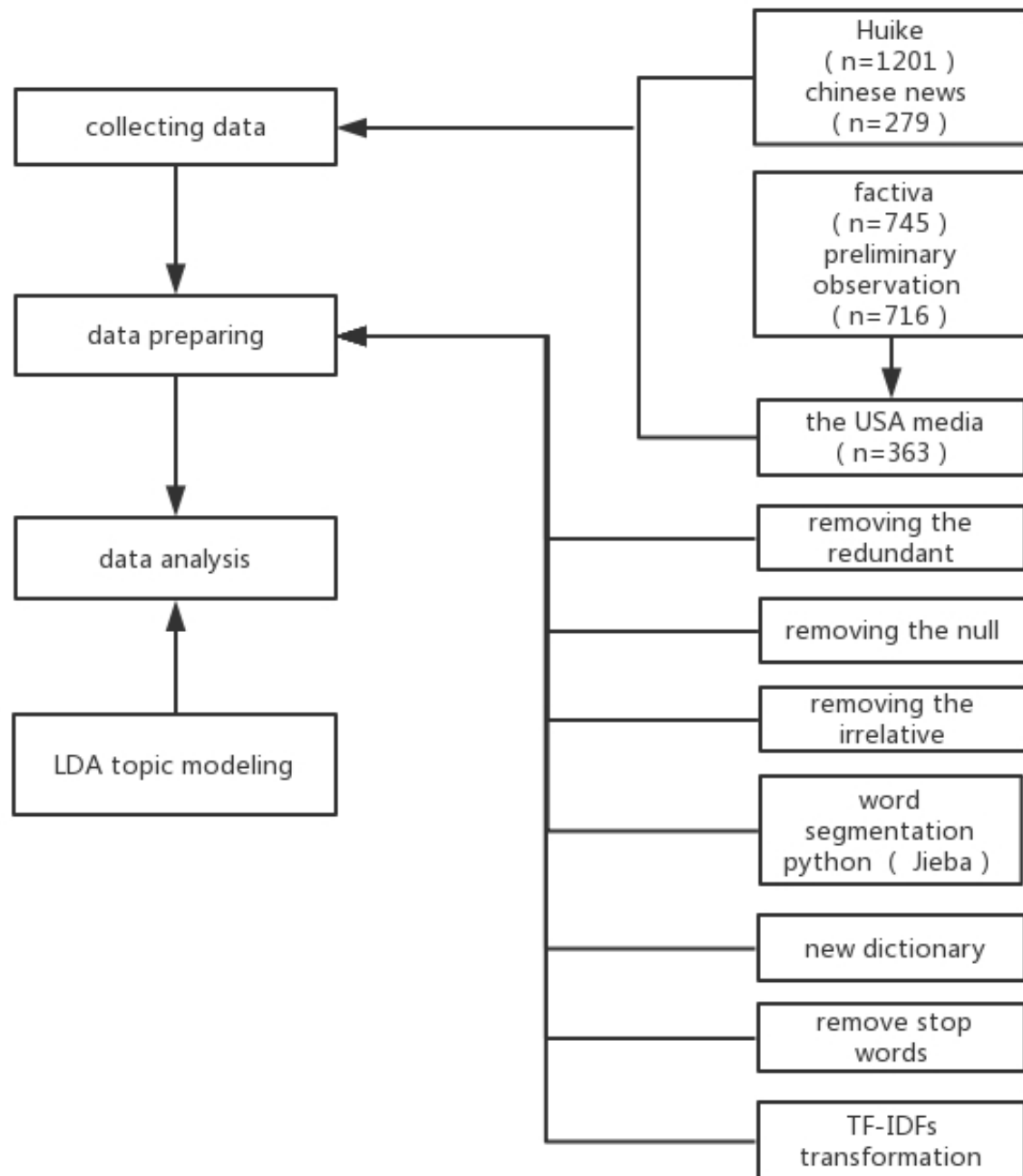


Table 1. Topic classification and keywords.

Classification group and topic name	Key words	Percentage (%)
Related diseases		
Topic 1: cancer	Lung cancer, cancer, tumor, treatment, patient, risk factor, pollution, professor, China, population, prevalence, diseases pattern, prevention of air pollution	31.9
Topic 5: risks of smoking	Smoker, movement, body, nicotine, content, quitting smoking, experts, symptom related to dead smokers	10.1
Topic 7: diseases induced by smoking	Asthma, citizen, hospital, doctor, patient, treatment, smoker, time, long-term, breath, chairman, cause	4.9
Topic 3: susceptible population	Children, research, food, contact, cause, influence, environment, increase, body, clothes, smog, content, reveal, professor, indoor, smoker, female	11.4
Topic 4: quitting smoking	Quit smoking, smoker, smoke, hospital, drug, breath, doctor, work, smoker, content, one kind, introduction, treatment, chairman	11.1
Topic 8: relevant research	Introduction, reveal, cigarette, children, smoke, officer, smoker, increase, factor, place, reason, environment, relevant	3.2
Air and PM_{2.5}		
Topic 6: air quality	PM _{2.5} , indoor, concentration, severe, microgram, air, pollution, smog, influence, kitchen, cooking fume	9.7
Control and restrictions		
Topic 9: classic smoking control case	Shenzhen, tobacco control, activity, citizen, place, rule, investigation, smoker, work, over, condition, patients, quit smoking, control, indoor, public place, increase	2.6
Topic 10: public control	Public place, quit smoking, ban tobacco, rule, place, professor children, indoor, China, tobacco control, body, influence, body, female worker, reveal	0.1
Topic 2: control of quitting smoking	Quit smoking, ban tobacco, third-hand smoke, public place, place, rule, ban, indoor, control, work, country, smoke, society, China, Beijing, relevant, smoker, outdoor	15

Figure 2. Intertopic distance map.

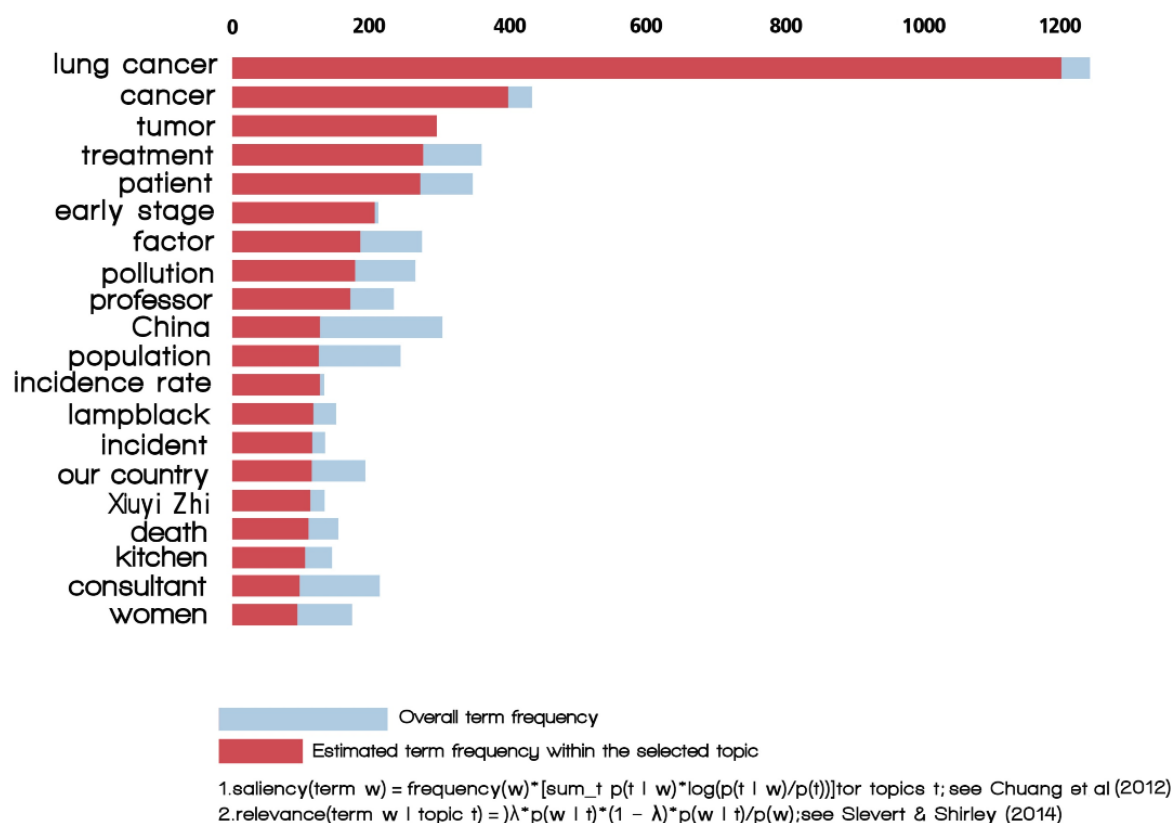
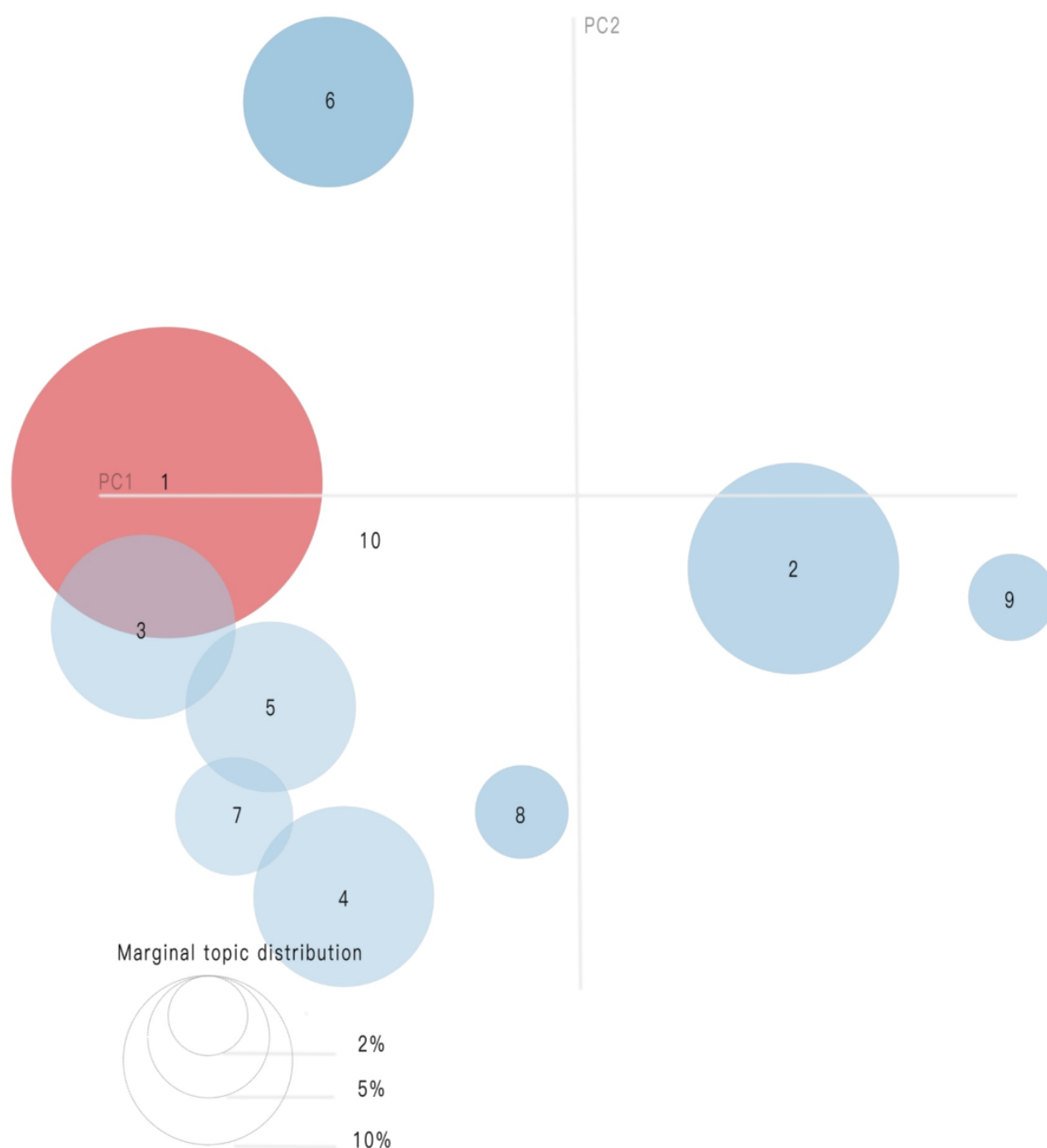


Figure 3. Top 20 most relevant terms for topic 1.

Results

We found 745 articles dated between January 1, 2013, and December 31, 2017. After excluding repeated articles, 716 news stories remained. This study collected 1201 news articles via the Wisser database platform. After culling repeated news and irrelevant entries, we recovered 288 news stories by selecting publications from only the largest media outlets in China representing the most influential mass media outlets. Finally, after the data-cleaning process was complete, we amassed 279 articles dated from January 1, 2013, through December 31, 2017. A comparison of the Chinese news and American news is listed in Table 2. Worldwide attention has been focused on THS [1,21], and the top five regions of origin of these articles

and their corresponding prevalence are listed in Table 3. Not surprisingly, the United States ranks first, publishing roughly 152 articles on THS from 2013-2017. There has been an obvious decrease in interest in THS from the Chinese media within the last 5 years.

Figure 2 presents an overall view of the topic model. We plotted different topics as circles, where overall prevalence was calculated as the areas of the circles. The centers of each topic were determined by computing the distance between topics; we used multidimensional scaling to represent the intertopic distances on a two-dimensional plane [22]. PC1 indicates the transverse axis and PC2 indicates the longitudinal axis in Figure 2.

Figure 3 shows a bar chart in a descending order of the top 20 most useful terms, for interpreting a topic. The overlaid bars represent a given term's corpus-wide frequency and the topic-specific frequency, as noted previously in the literature [23].

Figure 4 shows the percentage of news related to THS in China. Topic 1 (Cancer) was the most popular. Roughly 32% of all news stories noted this topic. Topic 2 (Control of quitting smoking) was included in approximately 15% news of THS-related news. Topic 3 (Susceptible population), Topic 4 (Quitting smoking), Topic 5 (Risks of smoking), and Topic 6 (Air quality) were each involved in roughly 10% of news stories. There were no news reports related to Topic 10.

Our results show that the Chinese press was less concerned with THS than the American press. Relative to 2013-2015, the number of reports in 2016 mentioning THS declined slightly. This trend indicates that the popularity of the topic might follow a worldwide trend (Table 3). We also observed a trend of increasing concern after a Chinese American scientist presented important findings about THS, linking THS to possible DNA damage that can cause diseases such as cancer [24,25]. This finding was widely reported both in China and the United States. It had an influence, particularly in government policy making (eg, legislation on smoke-free environments) [26]. Thus, there is a link between academic concerns and mass media concerns in China and the United States [5].

We also analyzed the three classifications, including links to diseases category, air and the PM_{2.5} category, and the control

and restrictions category (Table 1 and Figure 4). The first category, links to diseases, appeared in 72.6% of articles. This finding suggests that this category was frequently reported in China. Some health information can be misleading when delivered via the media, and new media platforms enable the rapid spread of news, some of it emotional or personal. This can include information that is false or misleading. As a result, the public still has a limited or even inaccurate understanding of THS. This situation was consistent with the findings of a prior study [27].

Roughly one tenth (9.7%) of the all-news stories were related to the air and PM_{2.5} category. The media seemed to focus on sensational topics related to THS, even if some links were tenuous. Therefore, professionalism and credibility in reporting are of vital importance. Furthermore, the control and restrictions category was included in 17.1% of all stories. This category is a unique one. While the Chinese media emphasizes control and restrictions, the American media focuses on conveying to smokers that they should quit the habit.

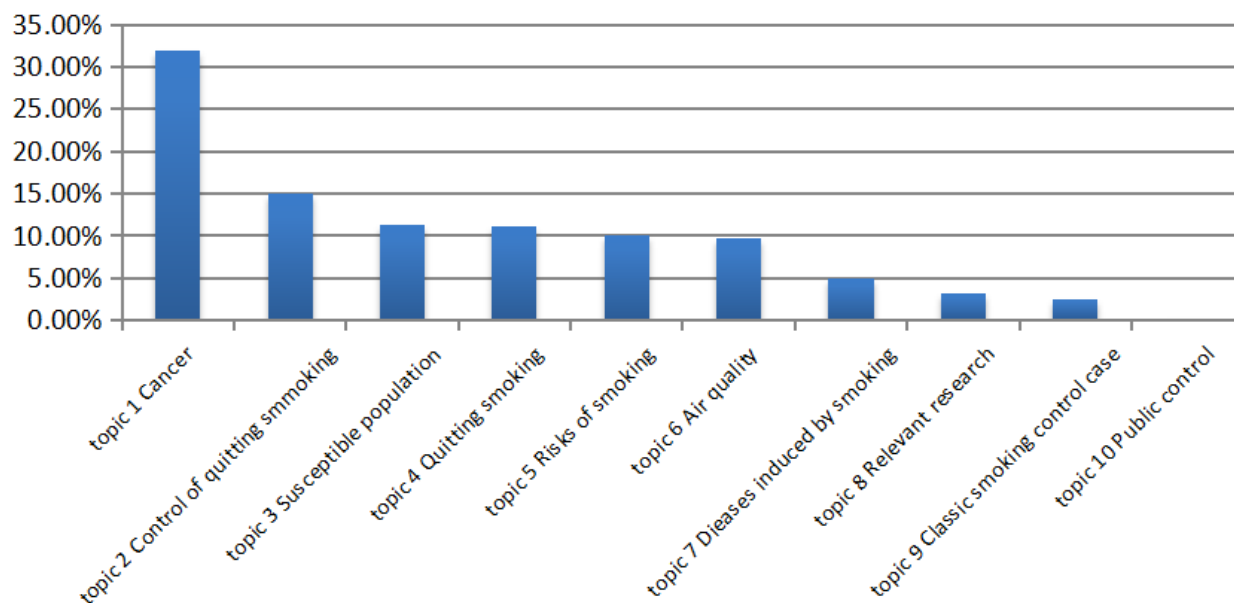
In comparing the Chinese and American news, we found an enormous variation between reports on the same topic in terms of the following three aspects. The Chinese news focused on children and the elderly as victims, whereas the American news focused more on pregnant women. The Chinese media discussed a number of measures and restrictions the government or authorities took or should be taken related to smoking, while the American press focused more on how individuals could eliminate risks. Furthermore, Chinese news articles were, on average, shorter than American news articles.

Table 2. Chinese and American news about thirdhand smoke.

Year	Chinese news articles (n)	American news articles (n)
2013	78	52
2014	57	114
2015	74	54
2016	29	38
2017	41	105
Total	279	363

Table 3. Countries publishing news about thirdhand smoke.

Country	Articles (n)
United States	152
United Kingdom	56
Canada	37
India	33
Australia	21

Figure 4. Topic percentage allocation.

Discussion

Principal Findings

Thirdhand smoke has been a growing topic for years in China. Topic modeling offers insights into understanding news reports related to THS. This study collected media reports from the United States and China and analyzed them using topic modeling. Specifically, we retrieved Chinese reports about THS from major newspaper and press outlets in China and compared this information with English-language news on the same subject from the United States. Our results revealed that the American press was more concerned with THS than the Chinese. In the Chinese media, three major concerns emerged: links of THS to disease, air and PM_{2.5}, and tobacco control and restrictions.

Our results indicate that the media has served a public health function by publishing reports that warn the general public about THS-related dangers such as lung cancer, asthma, tumors, and other diseases linked to tobacco. Furthermore, these articles emphasize the possible risks of THS and the susceptible populations (eg, children) in indoor environments. Smoking cessation and relevant research are also reported. Therefore, the media do communicate the risks associated with smoking, as well as information about prevention and smoking cessation.

For air quality and PM_{2.5}-related reports, the media attempts to connect sensational concepts with the topic, even if such concepts are not quite related. Therefore, professionalism and credibility are of vital importance. In addition, new media platforms enable the rapid spread of news that is more emotional or personal, or possibly false, which is a serious concern. Aided by new media platforms such as WeChat, Weibo, and Jinri Toutiao (the biggest new media platforms in China), information

that may include misleading or exaggerated concepts can quickly be disseminated on personalized newsfeeds.

In terms of topics related to tobacco control and restrictions, the Chinese media emphasizes control and restrictions more than the United States. The American media focuses on helping smokers quit the habit.

Strengths and Limitations

Topic modeling is a new method that reveals the major topics in media reports and singles out several key concerns and findings related to the topics. Data analysis and the visualization of news articles can generate useful information. However, there is a limitation to be noted in our study. We included only major media databases, which might omit some news content from new media, such as WeChat posts. Therefore, we may have missed some news stories.

Conclusion

Thirdhand smoke is an important public health issue. Collecting media reports on THS from major media outlets and analyzing this subject using topic modeling can facilitate a better understanding of the role that the media plays in communicating this health issue to the public. We conclude that the Chinese press still needs to consider fuller reports on THS rather than simply reporting sensational headlines and needs to show more professionalism by not publishing articles that lack scientific evidence. We recommend that additional studies be conducted related to sentiment analysis of news data to verify and measure the influence of topics revealed from the reports. For example, scientists could measure the educational function of the media for public health or study the influence of misleading information about THS generated by news reports.

Acknowledgments

This study was funded by the National Social Science Foundation of China (18CXW021).

Conflicts of Interest

None declared.

References

1. Winickoff JP, Friebely J, Tanski SE, Sherrod C, Matt GE, Hovell MF, et al. Beliefs about the health effects of "thirdhand" smoke and home smoking bans. *Pediatrics* 2009 Jan;123(1):e74-e79 [FREE Full text] [doi: [10.1542/peds.2008-2184](https://doi.org/10.1542/peds.2008-2184)] [Medline: [19117850](https://pubmed.ncbi.nlm.nih.gov/19117850/)]
2. Burton A. Does the smoke ever really clear? Thirdhand smoke exposure raises new concerns. *Environ Health Perspect* 2011 Feb;119(2):A70-A74 [FREE Full text] [doi: [10.1289/ehp.119-a70](https://doi.org/10.1289/ehp.119-a70)] [Medline: [21285011](https://pubmed.ncbi.nlm.nih.gov/21285011/)]
3. Protano C, Vitali M. The new danger of thirdhand smoke: why passive smoking does not stop at secondhand smoke. *Environ Health Perspect* 2011 Oct;119(10):A422 [FREE Full text] [doi: [10.1289/ehp.1103956](https://doi.org/10.1289/ehp.1103956)] [Medline: [21968336](https://pubmed.ncbi.nlm.nih.gov/21968336/)]
4. Quinlan K, McCaul K. Matched and mismatched interventions with young adult smokers: testing a stage theory. *Health Psychol* 2000 Mar;19(2):165-171. [Medline: [10762100](https://pubmed.ncbi.nlm.nih.gov/10762100/)]
5. Delgado-Rendon A, Cruz T, Soto D, Baezconde-Garbanati L, Unger J. Second and Thirdhand Smoke Exposure, Attitudes and Protective Practices: Results from a Survey of Hispanic Residents in Multi-unit Housing. *J Immigr Minor Health* 2017 Oct;19(5):1148-1155. [doi: [10.1007/s10903-016-0540-x](https://doi.org/10.1007/s10903-016-0540-x)] [Medline: [28074306](https://pubmed.ncbi.nlm.nih.gov/28074306/)]
6. Escoffery C, Bundy L, Carvalho M, Yembra D, Haardörfer R, Berg C, et al. Third-hand smoke as a potential intervention message for promoting smoke-free homes in low-income communities. *Health Educ Res* 2013 Oct;28(5):923-930 [FREE Full text] [doi: [10.1093/her/cyt056](https://doi.org/10.1093/her/cyt056)] [Medline: [23669213](https://pubmed.ncbi.nlm.nih.gov/23669213/)]
7. Zheng Y, Zhang Y, Larochelle H. A Deep and Autoregressive Approach for Topic Modeling of Multimodal Data. *IEEE Trans Pattern Anal Mach Intell* 2016 Dec;38(6):1056-1069. [doi: [10.1109/TPAMI.2015.2476802](https://doi.org/10.1109/TPAMI.2015.2476802)] [Medline: [26372202](https://pubmed.ncbi.nlm.nih.gov/26372202/)]
8. McLaurin E, McDonald A, Lee J, Aksan N, Dawson J, Tippin J, et al. Variations on a theme: Topic modeling of naturalistic driving data. *Proc Hum Factors Ergon Soc Annu Meet* 2014 Dec;58(1):2107-2111 [FREE Full text] [doi: [10.1177/1541931214581443](https://doi.org/10.1177/1541931214581443)] [Medline: [26190948](https://pubmed.ncbi.nlm.nih.gov/26190948/)]
9. Zhao W, Chen J, Perkins R, Liu Z, Ge W, Ding Y, et al. A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics* 2015;16 Suppl 13:S8 [FREE Full text] [doi: [10.1186/1471-2105-16-S13-S8](https://doi.org/10.1186/1471-2105-16-S13-S8)] [Medline: [26424364](https://pubmed.ncbi.nlm.nih.gov/26424364/)]
10. Hassanpour S, Langlotz CP. Unsupervised Topic Modeling in a Large Free Text Radiology Report Repository. *J Digit Imaging* 2016 Feb;29(1):59-62 [FREE Full text] [doi: [10.1007/s10278-015-9823-3](https://doi.org/10.1007/s10278-015-9823-3)] [Medline: [26353748](https://pubmed.ncbi.nlm.nih.gov/26353748/)]
11. Kandula S, Curtis D, Hill B, Zeng-Treitler Q. Use of topic modeling for recommending relevant education material to diabetic patients. *AMIA Annu Symp Proc* 2011;2011:674-682 [FREE Full text] [Medline: [22195123](https://pubmed.ncbi.nlm.nih.gov/22195123/)]
12. Goyal N, Gomeni R. A latent variable approach in simultaneous modeling of longitudinal and dropout data in schizophrenia trials. *Eur Neuropsychopharmacol* 2013 Nov;23(11):1570-1576. [doi: [10.1016/j.euroneuro.2013.03.004](https://doi.org/10.1016/j.euroneuro.2013.03.004)] [Medline: [23602612](https://pubmed.ncbi.nlm.nih.gov/23602612/)]
13. Yu K, Zhang J, Chen M, Xu X, Suzuki A, Ilic K, et al. Mining hidden knowledge for drug safety assessment: topic modeling of LiverTox as a case study. *BMC Bioinformatics* 2014;15 Suppl 17:S6 [FREE Full text] [doi: [10.1186/1471-2105-15-S17-S6](https://doi.org/10.1186/1471-2105-15-S17-S6)] [Medline: [25559675](https://pubmed.ncbi.nlm.nih.gov/25559675/)]
14. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. *Med Decis Making* 2014 Dec;34(3):343-351. [doi: [10.1177/0272989X13497998](https://doi.org/10.1177/0272989X13497998)] [Medline: [23901052](https://pubmed.ncbi.nlm.nih.gov/23901052/)]
15. Li A, Huang X, Hao B, O'Dea B, Christensen H, Zhu T. Attitudes towards suicide attempts broadcast on social media: an exploratory study of Chinese microblogs. *PeerJ* 2015;3:e1209 [FREE Full text] [doi: [10.7717/peerj.1209](https://doi.org/10.7717/peerj.1209)] [Medline: [26380801](https://pubmed.ncbi.nlm.nih.gov/26380801/)]
16. Blei D, Ng A, Jordan M. Latent Dirichllocation. *J Mach Learn Res* 2003 May 15;3:993-1022.
17. He B, De Sa C, Mitliagkas I, Ré C. Scan Order in Gibbs Sampling: Models in Which it Matters and Bounds on How Much. *Adv Neural Inf Process Syst* 2016;29 [FREE Full text] [Medline: [28344429](https://pubmed.ncbi.nlm.nih.gov/28344429/)]
18. Peng K, Liou L, Chang C, Lee D. Predicting Personality Traits of Chinese Users Based on Facebook Wall Posts. 2015 Oct 23 Presented at: 24th Wireless and Optical Communication Conference (WOCC); October 23-24, 2015; Taipei, Taiwan.
19. Adams B, McKenzie G. Crowdsourcing the character of a place: Character-level convolutional networks for multilingual geographic text classification. *Transactions in GIS* 2018 Jan 29;22(2):394-408. [doi: [10.1111/tgis.12317](https://doi.org/10.1111/tgis.12317)]
20. Grimmer J, Stewart B. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Polit. anal* 2017 Jan 4;21(03):267-297. [doi: [10.1093/pan/mps028](https://doi.org/10.1093/pan/mps028)]
21. Holitzki H, Dowsett L, Spackman E, Noseworthy T, Clement F. Health effects of exposure to second- and third-hand marijuana smoke: a systematic review. *CMAJ Open* 2017 Nov 24;5(4):E814-E822 [FREE Full text] [doi: [10.9778/cmajo.20170112](https://doi.org/10.9778/cmajo.20170112)] [Medline: [29192095](https://pubmed.ncbi.nlm.nih.gov/29192095/)]
22. Chuang J, Ramage D, Manning C, Heer J. Interpretation and trust: designing model-driven visualizations for text analysis. 2012 Oct 05 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 05-10, 2012; Austin, Texas p. A-52.

23. Chuang J, Manning C, Heer J. Termite: Visualization Techniques for Assessing Textual Topic Models. 2012 May 21 Presented at: Proceedings of the International Working Conference on Advanced Visual Interfaces; 2012; Capri Island, Italy p. 74-77.
24. Hang B, Sarker A, Havel C, Saha S, Hazra TK, Schick S, et al. Thirdhand smoke causes DNA damage in human cells. *Mutagenesis* 2013 Jul;28(4):381-391 [FREE Full text] [doi: [10.1093/mutage/get013](https://doi.org/10.1093/mutage/get013)] [Medline: [23462851](https://pubmed.ncbi.nlm.nih.gov/23462851/)]
25. Chen Y, Adhami N, Martins-Green M. Biological markers of harm can be detected in mice exposed for two months to low doses of Third Hand Smoke under conditions that mimic human exposure. *Food Chem Toxicol* 2018 Dec;122:95-103. [doi: [10.1016/j.fct.2018.09.048](https://doi.org/10.1016/j.fct.2018.09.048)] [Medline: [30253244](https://pubmed.ncbi.nlm.nih.gov/30253244/)]
26. Kennedy R, Ellens-Clark S, Nagge L, Douglas O, Madill C, Kaufman P. A Smoke-Free Community Housing Policy: Changes in Reported Smoking Behaviour-Findings from Waterloo Region, Canada. *J Community Health* 2015 Dec;40(6):1207-1215. [doi: [10.1007/s10900-015-0050-0](https://doi.org/10.1007/s10900-015-0050-0)] [Medline: [26070870](https://pubmed.ncbi.nlm.nih.gov/26070870/)]
27. Acuff L, Fristoe K, Hamblen J, Smith M, Chen J. Third-Hand Smoke: Old Smoke, New Concerns. *J Community Health* 2016 Dec;41(3):680-687. [doi: [10.1007/s10900-015-0114-1](https://doi.org/10.1007/s10900-015-0114-1)] [Medline: [26512014](https://pubmed.ncbi.nlm.nih.gov/26512014/)]

Abbreviations

LDA: Latent Dirichlet Allocation

TF-IDF: frequency-inverse document frequency

THS: thirdhand smoke

PM: particulate matter

Edited by G Eysenbach; submitted 06.10.18; peer-reviewed by C Zhang, A Majmundar; comments to author 25.10.18; revised version received 14.12.18; accepted 05.01.19; published 29.01.19.

Please cite as:

Liu Q, Chen Q, Shen J, Wu H, Sun Y, Ming WK

Data Analysis and Visualization of Newspaper Articles on Thirdhand Smoke: A Topic Modeling Approach

JMIR Med Inform 2019;7(1):e12414

URL: <http://medinform.jmir.org/2019/1/e12414/>

doi: [10.2196/12414](https://doi.org/10.2196/12414)

PMID: [30694199](https://pubmed.ncbi.nlm.nih.gov/30694199/)

©Qian Liu, Qiuyi Chen, Jiayi Shen, Huailiang Wu, Yimeng Sun, Wai-Kit Ming. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 29.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Detection of Bleeding Events in Electronic Health Record Notes Using Convolutional Neural Network Models Enhanced With Recurrent Neural Network Autoencoders: Deep Learning Approach

Rumeng Li¹, MS; Baotian Hu², PhD; Feifan Liu³, PhD; Weisong Liu², PhD; Francesca Cunningham⁴, PharmD; David D McManus^{3,5}, MSc, MD; Hong Yu^{1,2,3,6}, PhD

¹College of Information and Computer Science, University of Massachusetts Amherst, Amherst, MA, United States

²Department of Computer Science, University of Massachusetts Lowell, Lowell, MA, United States

³Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

⁴Department of Veterans Affairs, Center for Medication Safety, Hines, IL, United States

⁵Cardiology Division, Department of Medicine, University of Massachusetts Medical School, Worcester, MA, United States

⁶Center for Healthcare Organization and Implementation Research, Bedford Veterans Affairs Medical Center, Bedford, MA, United States

Corresponding Author:

Hong Yu, PhD

Department of Computer Science

University of Massachusetts Lowell

1 University Avenue

Lowell, MA, 01854

United States

Phone: 1 9789343620

Fax: 1 9789343551

Email: hong_yu@uml.edu

Abstract

Background: Bleeding events are common and critical and may cause significant morbidity and mortality. High incidences of bleeding events are associated with cardiovascular disease in patients on anticoagulant therapy. Prompt and accurate detection of bleeding events is essential to prevent serious consequences. As bleeding events are often described in clinical notes, automatic detection of bleeding events from electronic health record (EHR) notes may improve drug-safety surveillance and pharmacovigilance.

Objective: We aimed to develop a natural language processing (NLP) system to automatically classify whether an EHR note sentence contains a bleeding event.

Methods: We expertly annotated 878 EHR notes (76,577 sentences and 562,630 word-tokens) to identify bleeding events at the sentence level. This annotated corpus was used to train and validate our NLP systems. We developed an innovative hybrid convolutional neural network (CNN) and long short-term memory (LSTM) autoencoder (HCLA) model that integrates a CNN architecture with a bidirectional LSTM (BiLSTM) autoencoder model to leverage large unlabeled EHR data.

Results: HCLA achieved the best area under the receiver operating characteristic curve (0.957) and F1 score (0.938) to identify whether a sentence contains a bleeding event, thereby surpassing the strong baseline support vector machines and other CNN and autoencoder models.

Conclusions: By incorporating a supervised CNN model and a pretrained unsupervised BiLSTM autoencoder, the HCLA achieved high performance in detecting bleeding events.

(*JMIR Med Inform* 2019;7(1):e10788) doi:[10.2196/10788](https://doi.org/10.2196/10788)

KEYWORDS

autoencoder; BiLSTM; bleeding; convolutional neural networks; electronic health record

Introduction

Background and Significance

Bleeding is defined as the escape of blood from the circulatory system (arteries and veins) due to trauma, anatomic malformation, bleeding disorder, medications, and aging. Bleeding events include symptoms like reddening or darkening of urine or stools, bleeding of gums, blood blisters, bruises, and vomiting of blood. Studies show that high incidences of bleeding events are associated with cardiovascular disease in patients on anticoagulant therapy [1-9], which has contributed to its standing as the most-frequent adverse drug events (ADEs) [1,3-9]. Anticoagulants are considered a high-alert medication by the Institute for Safe Medication Practices because of the potential severity of anticoagulant-related bleeding. In a study on patients receiving oral anticoagulant therapy, major bleeding occurred at a rate of 7.22 per 100 person-years and fatal bleeding occurred at a rate of 1.31 per 100 person-years, with a case-fatality rate of 13.4% for major bleeding [3]. Adverse health outcomes resulting from bleeding include poor functional status, myocardial infarction, heart failure, stroke, and even death [3-9]. Prompt and accurate detection of bleeding events is essential to prevent such adverse health outcomes and improve drug-safety surveillance and pharmacovigilance.

Bleeding events are frequently not recorded in the structured fields and are buried in the electronic health record (EHR) notes [10]. Manual abstraction is prohibitively expensive. Rapid, accurate, and automated detection of bleeding events in EHR notes may have significant cost and logistical benefits over manual detection. Therefore, this study aimed to develop natural language processing (NLP) approaches to automatically detect bleeding events in EHR notes.

NLP approaches have demonstrated increasing utility in clinical text mining in recent years [11-14]. Deep neural network methods have recently achieved new state-of-the-art performance in a wide range of NLP tasks [15-17]. In this study, we explored deep learning models and compared them with the strong traditional machine-learning classifiers (eg, support vector machines [SVM]).

Two architectures of deep neural networks relevant to this work include convolutional neural network (CNN) [18] and recurrent neural network (RNN) with its variants of long short-term memory (LSTM) [19] and gated recurrent unit [20]. Both architectures have demonstrated advantages in text-processing tasks. The CNN models use layers with convolutional filters that are applied to local features [18] and therefore are able to capture local relationships between neighboring w -gram words in a sentence, but are less efficient for long-distance dependencies. In contrast, the LSTM models [19] are designed to learn long-term dependencies by maintaining an internal state, which represents the memory cell of the LSTM neuron. Thus, the LSTM models are able to memorize information for a longer duration than the CNN models. Bleeding events can be inferred by local context. Therefore, we chose CNN as the major model of our architecture, but leveraged the LSTM model to learn sentence-level representation. Our CNN model differs from the previous neural network models in that we deployed

an autoencoder neural network [21] as an unsupervised learning algorithm to learn a latent representation from unlabeled sentences in order to help improve CNN performance. Specifically, we propose the hybrid CNN and LSTM (HCLA) autoencoder model, which employs a CNN model that is integrated with a bidirectional LSTM (BiLSTM)-based autoencoder model to classify whether a sentence contains a bleeding event.

The knowledge-acquisition bottleneck problem presents a unique challenge in clinical NLP. Unlike data collection in the open domain, crowdsourcing methods (eg, Amazon Mechanical Turks [22]) cannot be easily applied to medical domain data collection due to privacy concerns. Annotation by medical professionals is expensive and time consuming, and annotated data in the clinical domain are typically limited. Because our HCLA model leverages a large number of unlabeled EHR notes, our results demonstrate that domain-specific features learned through such an autoencoder can effectively improve the supervised learning performance, despite the small amount of the training data.

Related Works

Existing work in automated bleeding detection mainly involves detection and classification of bleeding for wireless capsule endoscopy images. Neural network methods are also employed for such image detection [23,24]. In addition, previous studies have assessed detection of bleeding events in outcome studies by using health registers [25]. However, studies on the detection of bleeding events in EHR notes are lacking.

The proposed model is based on neural network models that learn feature representations for sentence-level classification. Related work includes the CNN models that first made a series of breakthroughs in the computer vision field and subsequently showed excellent performance in NLP tasks such as machine translation [26], sentence classification [27,28], and sentence modelling [29]. Autoencoders were originally proposed to reduce the dimensionality of images and documents [21] and were subsequently applied to many NLP tasks such as sentiment analysis [30], machine translation [31], and paraphrase detection [32].

Neural network models have been applied to the clinical data-mining tasks. Gehrmann et al [33] applied CNNs to 10 phenotyping tasks and showed that they outperformed concept extraction-based methods in almost all tasks. CNN was used to classify radiology free-text reports and showed an accuracy equivalent to or beyond that of an existing traditional NLP model [34]. Lin et al [35] also used a CNN model to identify the International Classification of Diseases, Tenth Revision, Clinical Modification, diagnosis codes in discharge notes and showed outstanding performance compared with traditional methods; they also showed that the convolutional layers of the CNN can effectively identify keywords for use in the prediction of diagnosis codes. Since our annotated data are relatively small, we expanded the CNN model by integrating it with an LSTM-based autoencoder. Tran et al [36] developed two independent deep neural network models: one based on CNNs and the other based on RNNs with hierarchical attention for the prediction of mental conditions in psychiatric notes; their study showed that the CNN and RNN models outperformed the

competitive baseline approaches. Furthermore, a previous study used semisupervising learning methods such as learning from positive and unlabeled examples [37] and the anchor-and-learn method [38], for which traditional machine-learning algorithms like expectation-maximization and SVM can be used to build classifiers.

Methods

Models

The EHR notes used in this study were provided by the University of Massachusetts Medical School. Our study was approved by the Institutional Review Board. All the EHR notes were deidentified.

Our HCLA model integrates a supervised CNN architecture and an unsupervised BiLSTM-based autoencoder. Given a sentence as input, we trained a classifier to determine whether the sentence contains a bleeding event. Figure 1 illustrates the architecture of our model. After passing the sentence through the CNN model and the BiLSTM-based autoencoder simultaneously, the HCLA model generates two separate representations for the input sentence: local features encoded by the CNN model and global features encoded by the autoencoder. A softmax function was then used to determine whether the input contained a bleeding event.

Bidirectional Long Short-Term Memory-Based Autoencoder

An autoencoder is a neural network that typically has three layers: an input layer, a hidden (encoding) layer, and a decoding layer. Through the encoding process, the inputs are compressed into a hidden representation, which is then used to reconstruct the input back in the decoding process.

A BiLSTM-based autoencoder has two major parts: encoder and decoder. During the encoding phase, an LSTM is used to scan the input $X=\{x_1, x_2... x_i... x_n\}$ in a sequential order. Each time, it takes a word x_i with e_i as its embedding and the hidden representation h_{i-1} generated at the previous step as the input to generate the representation h_i for the current step.

The final is h_n the representation of the input sentence. A BiLSTM model uses the same LSTM to scan the input sentence again in reverse and obtains another representation h'_n , so that the input is encoded as the concatenated hidden representation $h_n, h'_n \in R^k$, where k is a predefined dimensionality.

The decoder is another LSTM layer. The hidden representation is fed to the decoder LSTM layer to reconstruct the input words. First, we set $h'_0=h_n$ to repeat the following steps until the input is reconstructed:

$$h'_i=LSTM(h'_{i-1}, e_{i-1})$$

$$o_i=Wh'_i+b$$

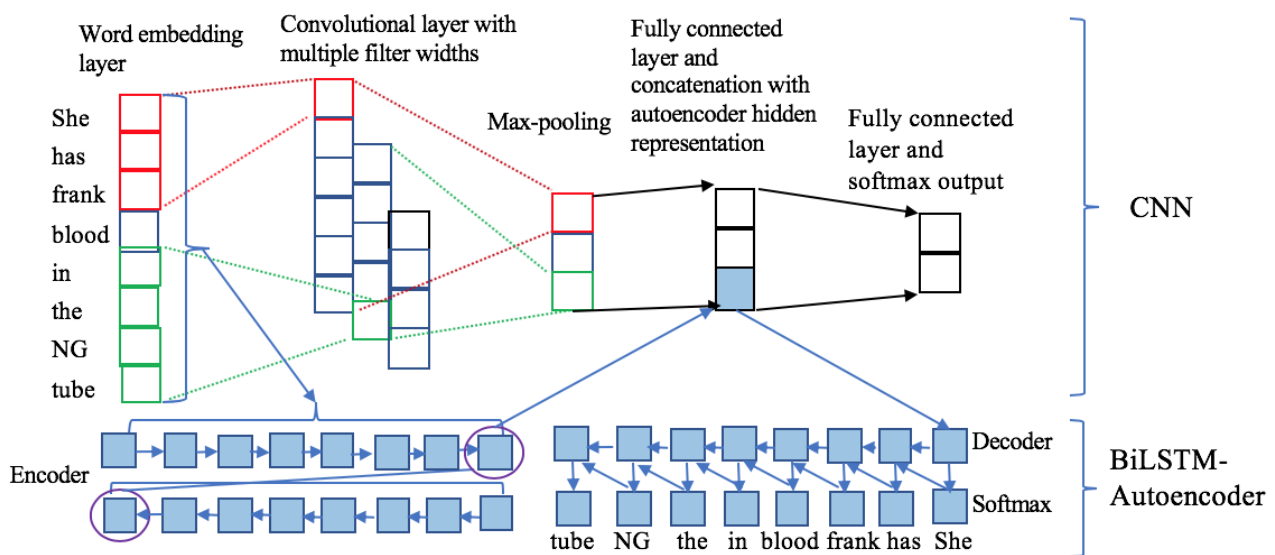


$$e_i = emb(x_i)$$

The LSTM takes the h'_{i-1} that is the hidden state of the previous step and e_{i-1} that is the word generated in the previous step as input and updates h'_{i-1} to h'_i . Subsequently, h'_i is passed through a softmax layer to generate the word at i^{th} step x_i .

After training by the abovementioned procedures, the hidden representation h_n is obtained as a condensed and better representation of the sentence.

Figure 1. The hybrid convolutional neural network (CNN) and long short-term memory (LSTM) autoencoder (HCLA) model architecture with two major components: the CNN model and the bidirectional LSTM (BiLSTM)-based autoencoder. NG: nasogastric.



Convolutional Neural Networks

The CNN model takes a sequence of words as input and outputs a fixed-length low-density vector as a representation of the

input. Words are first represented using their embeddings, which can be learned during training or loaded from pretrained models. We will report how we set the embeddings for each of our specific models in the Models subsection. The sentence

subsequently becomes a matrix whose dimensions are decided by the word numbers of the input and the dimensions of word embeddings. Convolutional layers of different sizes scan the matrix to generate a new and dense representation of the input. The newly generated representations are further projected to a fixed-length vector through max pooling as the final representation of the input. We adapted the architecture of an open-domain CNN model [27] with the following components:

1. Input layer: The input is a sequence of tokens $X=\{x_1, x_2, \dots, x_n\}$ where n denotes the length of the sentence. Token x_i is associated with a d -dimensional embedding e_i , $e_i=\{e_i^1, e_i^2, \dots, e_i^d\}$. Therefore, the input is represented as a feature map of dimensionality $n \times d$.
2. Convolutional layer: A convolutional operation with a filter sliding over the input is applied for local feature learning. Given the input sentence $X=\{x_1, x_2, \dots, x_n\}$ (zero padded if necessary), a feature c_i can be learned from a window of words $[x_{i+k-1}]$ as $c_i=f(w[x_{i+k-1}] + b)$, where w is the convolutional weights $w \in R^{d \times kd}$, bias $b \in R^d$, f is the nonlinear activation function (we used a hyperbolic tangent in our experiments), and k is the filter width. In our model, we used three filters of width 3, 4, and 5, and the number of filters used was 10.
3. Max-pooling layer: A max operation is applied to the result of each filter to keep the most-salient information and reduce dimensionality [27]. The outputs of the three filters after max pooling are concatenated in this framework.
4. Concatenation layer: After the max-pooling layer was fully connected, at this step, we concatenated the final hidden representation [27] of the encoder layer obtained from the BiLSTM-based autoencoder approach, as described in the above Bidirectional Long Short-Term Memory-Based Autoencoder subsection.
5. Softmax layer: Another fully connected layer and softmax operation was applied for the prediction. The cut-off point of 0.5 was used to convert the predicted probability to a binary outcome with regard to whether the sentence contains a bleeding event.

Results

Experimental Datasets

We expert annotated a corpus of 878 longitudinal EHR notes of patients with cardiovascular events. This corpus contains

Textbox 1. Sentence samples from our dataset with 2 positive bleeding samples and 2 negative bleeding samples. POS: positive bleeding sample sentence; NEG: negative bleeding sample sentence.

POS 1: Patient was admitted with hematemesis and blood per rectum.

POS 2: Anticoagulation has been held on this patient secondary to recent gastrointestinal bleed.

NEG 1: She has done well with the warfarin with no further thromboembolic episodes and no bleeding problems.

NEG 2: The patient is also on Keppra for seizure activity, and he has been seizure-free on that medication.

76,577 sentences and 562,630 word-tokens. Each note was annotated by at least two physicians. The interannotation agreement (Cohen k) among the annotators was $k=0.9182$.

We preprocessed the data and removed duplicate sentences. In addition, we removed sentences with length <5 word-tokens, as those were mostly incomplete sentences or sentence fragments. The remaining 48,628 sentences were used for training and validation. Textbox 1 presents the representative sentences in our annotated corpus.

Although we had a total of 878 annotated notes, only 291 notes mentioned bleeding events, from which we identified a total of 1451 sentences. These 1451 sentences were considered as “positive” data for training. From the 291 notes, a total of 285 sentences that mentioned bleeding events, but were annotated as negation. Those sentences presented as “harder” examples for our NLP systems, and we included them in the “negative” data for training.

We performed downsampling to include equal number of positive and negative data for training and validation. Since we had included 285 sentences that contain negated bleeding events, 1166 negative sentences were randomly selected from the remaining negative sentences in the 291 notes, as those sentences were more challenging for an NLP system to identify than sentences from the notes that did not contain a bleeding event. Although we did not include any sentence from the 587 annotated notes that did not contain a bleeding event, the 32,704 sentences from those notes were included to train the autoencoder on feature representation.

Of the combined 2902 positive and negative sentences (50% each), we randomly selected 1 of 10 sentences as the testing data, and the other 9 sentences were used for model training. By setting aside the testing data, we trained the BiLSTM-based autoencoder model on all the remaining sentences of the 878 notes, with a total of 48,338 sentences after preprocessing.

An NLP system that is trained for downsampling may not perform well for the real-world data; in our case, the positive and negative data were highly unbalanced (only 2.9% of annotated sentences contained a bleeding event). To accurately evaluate the performance of our NLP system, we annotated 6 additional EHR notes as unseen hold-out data for testing. Those 6 notes included a total of 2,345 sentences, of which only 64 sentences (2.7%) were positive.

Experimental Setup

We implemented the neural network models in Pytorch [39]. For the evaluation metrics, we used precision, recall, and F-score and reported the overall performance as well as the performance on positive instances and negative instances. In addition, we reported the area under the receiver operating characteristic curve (AUC-ROC) score. The AUC-ROC plots the true positive rate (y-coordinate) against the false positive rate (x-coordinate) at various threshold settings. For the testing on the natural EHR notes, we reported overall accuracy as well as the precision, recall, and F-score on positive samples. All the word embedding sizes were initialized or pretrained with 200 dimensions. The dropout rate was set as 0.3, batch size as 16, and learning rate as 0.001. Optimization was performed using stochastic gradient descent. For the BiLSTM autoencoder model, the number of hidden neurons was 64.

Models

We conducted experiments with the following baselines to compare them with our proposed model. We first compared our model with a strong traditional machine-learning classifier as the SVM model. The basic bag-of-words features are used for SVM. A stronger SVM baseline with both bag-of-words features and bag of the unified medical language system (UMLS) [40] concept features were implemented for comparison. In addition, as the proposed model was an integrated model including a word-embedding CNN and a BiLSTM autoencoder, we conducted experiments to determine the separate performance of the two components. Pretrained word embeddings were used in additional experiments to examine their effect in model performance.


Support Vector Machines+Bag of Words

A standard linear SVM classifier [41] with bag-of-words features was used. Parameter C (penalty parameter of the error term) was set as 1, and other parameters were set as default in the sklearn.SVM.SVC implementation [41].

Support Vector Machines+Bag of Words+Unified Medical Language System

A standard linear SVM classifier [41] with bag-of-words features and bag of UMLS [40] concept features were used. Further, we used MetaMap [42], a tool created by NLM that maps from free text to biomedical concepts in the UMLS, to identify medical phrases. The same parameters were used as mentioned in the Support Vector Machines+Bag of Words subsection above.

Autoencoder

The BiLSTM-based autoencoder model has been described above (see Methods section); all word embedding was randomly initialized and modified during training. A fully connected layer was used to connect the obtained hidden representation of  followed by a softmax operation to generate the prediction.

Autoencoder and Pretrained Word Embedding

This model was similar to the autoencoder model. However, in this model, we pretrained the word embedding on 4.7 million

EHR notes using the Glove model [43]. The pretrained vectors were fine tuned for the task during training.

Convolutional Neural Network

The CNN model used has been described in the Methods section, with all word embeddings randomly initialized and modified during training.

Convolutional Neural Network and Pretrained Word Embedding

This model was similar to the CNN model. The same pretrained word embeddings described in the Autoencoder and Pretrained Word Embedding experiment were used, and the vectors were fine tuned for the task during training.

Hybrid Convolutional Neural Network and Long Short-Term Memory Autoencoder

Our proposed model incorporates CNN, pretrained word embedding, and a BiLSTM-based autoencoder, as described in the Methods section.

Convolutional Neural Network for Negation Bleeding

As negation bleeding (eg, NEG1 in [Textbox 1](#)) is a relatively difficult and misleading subset of the corpus for the model to make predictions, we conducted this extra experiment with only the 285 sentences that mentioned negated bleeding as negative samples. Of these, 185 sentences were used with the 1451 positive samples for training, and the remaining 100 sentences were used for testing.

Hybrid Convolutional Neural Network and Long Short-Term Memory Autoencoder for Negation Bleeding

The same data setting was used as mentioned above for Convolutional Neural Network for Negation Bleeding.

Experimental Results

Our HCLA model showed the best performance across all evaluation metrics, with an AUC-ROC value of 0.957, overall F-score of 0.938, and F-scores of 0.932 and 0.943 for positive and negative sentences, respectively ([Table 1](#)). With pretrained word embedding, both the autoencoder and the CNN models performed better than learning word representation directly from the data. For the traditional SVM model, improved performance was achieved by incorporating the UMLS knowledge, leading to an overall F-score of 0.886 and an AUC-ROC value of 0.934 compared to an F-score of 0.862 and an AUC-ROC of 0.921 without the UMLS features. The incorporation of UMLS knowledge especially improved the precision score on positive samples with a large increase of 0.043. The precision score of all CNN models demonstrated a consistent increase in all positive and negative samples. As shown in [Table 1](#), the CNN model without the autoencoder outperformed the model that was solely built on the autoencoder.

To further evaluate our model's performance on natural EHR notes (as compared to negative sampling), we tested the proposed HCLA model on the 6 extra annotated notes. The resulting overall accuracy outcomes and precision, recall, and F-score on positive samples are shown in [Table 2](#).

Table 1. Comparison of the study results at baseline.

Model	Positive			Negative			Overall			AUC-ROC ^a
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	
SVM ^b +BOW ^c	0.848	0.883	0.865	0.878	0.841	0.859	0.862	0.862	0.862	0.921
SVM+BOW+UMLS ^d concept	0.891	0.887	0.889	0.870	0.886	0.878	0.886	0.886	0.886	0.934
Autoencoder	0.861	0.855	0.858	0.856	0.862	0.859	0.859	0.859	0.859	0.920
Autoencoder+pretrained word embedding	0.875	0.869	0.872	0.870	0.876	0.873	0.872	0.872	0.872	0.926
CNN ^e	0.908	0.877	0.892	0.879	0.910	0.894	0.893	0.893	0.893	0.938
CNN+pretrained word embedding	0.930	0.911	0.920	0.912	0.931	0.921	0.921	0.921	0.921	0.946
HCLA ^f	0.954	0.912	0.932	0.925	0.961	0.943	0.938	0.938	0.938	0.957
CNN for negation bleeding	N/A ^g	N/A	N/A	0.820	0.820	0.820	N/A	N/A	N/A	N/A
HCLA for negation bleeding	N/A	N/A	N/A	0.860	0.860	0.860	N/A	N/A	N/A	N/A

^aAUC-ROC: area under the receiver operating characteristic curve.

^bSVM: support vector machines.

^cBOW: bag of words.

^dUMLS: unified medical language system.

^eCNN: convolutional neural network.

^fHCLA: hybrid convolutional neural network and long short-term memory autoencoder.

^gN/A: not applicable.

Table 2. Performance of the hybrid convolutional neural network and long short-term memory autoencoder model on natural electronic health record notes.

Performance parameter	Value
Overall accuracy	0.938
Precision on positive samples	0.992
Recall on positive samples	0.944
F-score on positive samples	0.967

Discussion

Principal Findings

This study addresses the detection of important bleeding events in EHR notes. Clinical phenotyping is challenging mainly due to irregularity of clinical narratives, which incorporates domain-specific medical jargon, abbreviations, incorrect use of natural language (eg, spelling errors), etc [44,45]. In addition, negation is common in the clinical domain, which adds complexity. The difficulty of the NLP task is further exacerbated by the limited size of human annotated gold standard positive samples, which makes it difficult for our data-driven deep learning approaches to extract effective features. However, the proposed HCLA model (Textbox 1) achieved the best result of 0.938 (F-score) and 0.957 (AUC-ROC value).

Our results show that all end-to-end CNN models outperformed the baseline SVM model, despite the incorporation of knowledge-based features for SVM. When we pretrained word embedding over large unlabeled EHR notes (4.3 million), the overall F-score performance improved by approximately 0.03, demonstrating the effectiveness of using unlabeled data. Our

results demonstrate that the BiLSTM-based autoencoder improved sentence representation. By concatenating the representation of the BiLSTM-based autoencoder, we further improved the performance by 0.017 (from 0.921 to 0.938) and the AUC-ROC value by 0.957, even though the model was trained with a relatively small data set (1451 positive samples and 285 negative samples).

Our annotated data incorporated both bleeding and negated bleeding events. Detection of bleeding signals is challenging from narratives, but separating true bleeding events from negated bleeding events is more challenging due to different negation variations. We therefore evaluated how our model performed in terms of accurately identifying negated bleeding events. Comparing the data setting described above for CNN and HCLA for negation bleeding, the CNN model achieved an accuracy of 0.82 and our proposed model achieved an accuracy of 0.86. The results validate the ability of the model to learn to grasp meaningful features in the dataset, rather than just depend on “bleeding”-related word indicators. For example, in Textbox 1, the NEG1 sentence contains the word “bleeding” but describes a nonbleeding event.

Textbox 2. Misclassified sentence samples by the proposed hybrid convolutional neural network and long short-term memory autoencoder system. NEG: negative bleeding sample sentence; POS: positive bleeding sample sentence.

NEG1: Here is no chest pain, chest pressure, shortness breath, wheezing, coughing, nausea, vomiting, or gastrointestinal bleeding.

POS1: At that time, he was found to be anemic with heme-positive stools and was admitted for evaluation.

POS2: There is no evidence or stigmata for recent bleed.

Table 2 presents the model's performance in real applications in the 6 natural EHR notes. In this unbalanced data of total 2345 sentences with only 64 positive sentences (2.7%), our model showed good results. The overall accuracy was 0.938. The model achieved high precision, recall, and F-score of 0.992, 0.944, and 0.967, respectively, on positive samples. Performance in positive samples is meaningful because it reflects how accurately the NLP system detects bleeding events in the EHR notes.

Error Analysis

Our error analyses showed that HCLA needs improvement in negation detection and analyzing complex sentences. **Textbox 2** includes representative examples to show that HCLA failed the classification. The first sentence in **Textbox 2** was a negated bleeding sample, but was misclassified as a positive instance. Since our employed CNN model mainly focused on the local context, it may fail to recognize the distant negation cue "no." The second sentence was a positive sample, but was misclassified as a negative sample by the system. The second sentence is complex and required knowledge inference, which may be challenging for the model.

The third sentence in **Textbox 2** was annotated as a positive instance, but seemed to be correctly identified by our system as a negative instance. On examining the note, we found that it was a follow-up of a patient whose chief complaint was hemorrhoidal bleeding. Our annotators annotated bleeding events at the whole-note level. Although the bleeding event in this specific sentence in the section of Physical Examination, seemed to be negative, it was annotated as "bleeding present." In this case, the annotators interpreted the bleeding as the present complaint, even though the sentence clearly indicated no evidence of a recent bleed.

This sentence represents a complex case of annotation consistency. The annotation guideline needs to be updated to

refine the definition of "assertion." On the other hand, this annotation highlights one limitation of our NLP system that is based on sentence-level classification. Our future work will focus on exploring the context of the whole note.

Study Strength

The contributions of this study are several folds: This study is the first to automatically detect bleeding events from EHR notes. We have demonstrated the effectiveness of HCLA as a high-performance bleeding event-detection NLP system from EHRs. In addition, we have demonstrated the effectiveness of the HCLA architecture that can be trained from small annotated data.

Limitations

We acknowledge a few limitations to this study. The gold dataset for our experiments was relatively small. Therefore, we built unsupervised models to leverage the large unlabeled EHR data in order to improve the performance. However, the generalizability and robustness of the model were not evaluated on a large scale. In addition, our system was based on sentence-level classification and does not consider the context of whole notes.

Conclusions

This is the first study to address bleeding detection in EHRs. Our proposed HCLA neural network model effectively outperformed standard CNN models, autoencoder models, and SVM models by using a limited number of expert annotations. In the future, we will attempt to use active learning methods in order to improve the efficiency of experts' annotation. Depending on more high-quality annotation, we will further mine data on bleeding causes, anatomic sites of bleeding, bleeding severity, and assertion (current vs history) from EHRs, as tasks are important and require further examination.

Acknowledgments

We wish to thank the UMassMed annotation team—WL, Steve Belknap, Nadya Frid, Alex Granillo, Heather Keating, and Victoria Wang—for creating the gold standard evaluation for this work. This study was supported, in part, by the grant R01HL12508 and R01HL125089 from the National Institutes of Health. Any opinions, findings, and conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect those of the sponsor.

Conflicts of Interest

None declared.

References

1. Shoeb M, Fang M. Assessing bleeding risk in patients taking anticoagulants. *J Thromb Thrombolysis* 2013 Apr;35(3):312-319 [[FREE Full text](#)] [doi: [10.1007/s11239-013-0899-7](https://doi.org/10.1007/s11239-013-0899-7)] [Medline: [23479259](https://pubmed.ncbi.nlm.nih.gov/23479259/)]

2. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation* 2017 Dec 07;135(10):e146-e603 [FREE Full text] [doi: [10.1161/CIR.0000000000000485](https://doi.org/10.1161/CIR.0000000000000485)] [Medline: [28122885](https://pubmed.ncbi.nlm.nih.gov/28122885/)]
3. Linkins L, Choi PT, Douketis JD. Clinical impact of bleeding in patients taking oral anticoagulant therapy for venous thromboembolism: a meta-analysis. *Ann Intern Med* 2003 Dec 02;139(11):893-900. [Medline: [14644891](https://pubmed.ncbi.nlm.nih.gov/14644891/)]
4. Wilcox C, Cryer B, Henk H, Zarotsky V, Zlateva G. Mortality associated with gastrointestinal bleeding events: Comparing short-term clinical outcomes of patients hospitalized for upper GI bleeding and acute myocardial infarction in a US managed care setting. *Clin Exp Gastroenterol* 2009;2:21-30 [FREE Full text] [Medline: [21694823](https://pubmed.ncbi.nlm.nih.gov/21694823/)]
5. Chai-Adisaksotha C, Crowther M, Isayama T, Lim W. The impact of bleeding complications in patients receiving target-specific oral anticoagulants: a systematic review and meta-analysis. *Blood* 2014 Oct 09;124(15):2450-2458 [FREE Full text] [doi: [10.1182/blood-2014-07-590323](https://doi.org/10.1182/blood-2014-07-590323)] [Medline: [25150296](https://pubmed.ncbi.nlm.nih.gov/25150296/)]
6. Davis SM, Broderick J, Hennerici M, Brun NC, Diringer MN, Mayer SA, Recombinant Activated Factor VII Intracerebral Hemorrhage Trial Investigators. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. *Neurology* 2006 Apr 25;66(8):1175-1181. [doi: [10.1212/01.wnl.0000208408.98482.99](https://doi.org/10.1212/01.wnl.0000208408.98482.99)] [Medline: [16636233](https://pubmed.ncbi.nlm.nih.gov/16636233/)]
7. Flibotte JJ, Hagan N, O'Donnell J, Greenberg SM, Rosand J. Warfarin, hematoma expansion, and outcome of intracerebral hemorrhage. *Neurology* 2004 Sep 28;63(6):1059-1064. [Medline: [15452298](https://pubmed.ncbi.nlm.nih.gov/15452298/)]
8. Oprea AD, Noto CJ, Halaszynski TM. Risk stratification, perioperative and periprocedural management of the patient receiving anticoagulant therapy. *J Clin Anesth* 2016 Nov;34:586-599. [doi: [10.1016/j.jclinane.2016.06.016](https://doi.org/10.1016/j.jclinane.2016.06.016)] [Medline: [27687455](https://pubmed.ncbi.nlm.nih.gov/27687455/)]
9. Wysowski DK, Nourjah P, Swartz L. Bleeding complications with warfarin use: a prevalent adverse effect resulting in regulatory action. *Arch Intern Med* 2007 Jul 09;167(13):1414-1419. [doi: [10.1001/archinte.167.13.1414](https://doi.org/10.1001/archinte.167.13.1414)] [Medline: [17620536](https://pubmed.ncbi.nlm.nih.gov/17620536/)]
10. Turchin A, Shubina M, Breydo E, Pendergrass M, Einbinder J. Comparison of information content of structured and narrative text data sources on the example of medication intensification. *J Am Med Inform Assoc* 2009;16(3):362-370 [FREE Full text] [doi: [10.1197/jamia.M2777](https://doi.org/10.1197/jamia.M2777)] [Medline: [19261947](https://pubmed.ncbi.nlm.nih.gov/19261947/)]
11. Shickel B, Tighe P, Bihorac A, Rashidi P. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE J Biomed Health Inform* 2018 Sep;22(5):1589-1604 [FREE Full text] [doi: [10.1109/JBHI.2017.2767063](https://doi.org/10.1109/JBHI.2017.2767063)] [Medline: [29989977](https://pubmed.ncbi.nlm.nih.gov/29989977/)]
12. Che Z, Cheng Y, Zhai S, Sun Z, Liu Y. Boosting Deep Learning Risk Prediction with Generative Adversarial Networks for Electronic Health Records. 2017 Presented at: IEEE International Conference on Data Mining (ICDM); 2017; New Orleans, USA p. 787 URL: <https://doi.org/10.1109/ICDM.2017.93> [doi: [10.1109/ICDM.2017.93](https://doi.org/10.1109/ICDM.2017.93)]
13. Rumeng L, Abhyuday NJ, Hong Y. A hybrid Neural Network Model for Joint Prediction of Presence and Period Assertions of Medical Events in Clinical Notes. *AMIA Annu Symp Proc* 2017;2017:1149-1158 [FREE Full text] [Medline: [29854183](https://pubmed.ncbi.nlm.nih.gov/29854183/)]
14. Jagannatha AN, Yu H. Bidirectional RNN for Medical Event Detection in Electronic Health Records. *Proc Conf* 2016 Jun;2016:473-482 [FREE Full text] [Medline: [27885364](https://pubmed.ncbi.nlm.nih.gov/27885364/)]
15. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, et al. arXivcs Internet. 2017 Jun 12. Attention Is All You Need URL: <https://arxiv.org/abs/1706.03762> [accessed 2018-12-17] [WebCite Cache ID 74k5OqeZI]
16. See A, Liu P, Manning C. Get To The Point: Summarization with Pointer-Generator Networks. 2017 Apr 14. URL: <https://arxiv.org/abs/1704.04368> [accessed 2018-12-17] [WebCite Cache ID 74k5S2yqd]
17. Hu M, Peng Y, Qiu X. Cornell University - arXivcs. 2017 May 8. Reinforced Mnemonic Reader for Machine Comprehension URL: <https://arxiv.org/abs/1705.02798> [accessed 2018-12-17] [WebCite Cache ID 74k5U5xhR]
18. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 1998;86:2278-2324 [FREE Full text]
19. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997 Nov 15;9(8):1735-1780. [Medline: [9377276](https://pubmed.ncbi.nlm.nih.gov/9377276/)]
20. Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.: Association for Computational Linguistics; 2014 Presented at: EMNLP; 2014; Doha, Qatar p. 1724-1734 URL: <http://arxiv.org/abs/1406.1078> [doi: [10.3115/v1/D14-1179](https://doi.org/10.3115/v1/D14-1179)]
21. Hinton G, Salakhutdinov R. Reducing the dimensionality of data with neural networks. *Science* 2006 Jul 28;313(5786):504-507 [FREE Full text] [doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647)] [Medline: [16873662](https://pubmed.ncbi.nlm.nih.gov/16873662/)]
22. Mturk. Amazon Mechanical Turk URL: <https://www.mturk.com/> [accessed 2019-01-14] [WebCite Cache ID 75QFGhNWX]
23. Jia X, Meng M. A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images. In: *Conf Proc IEEE Eng Med Biol Soc*. 2016 Dec Presented at: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).?642; 2016; Orlando, FL, USA p. 639-642 URL: <https://doi.org/10.1109/EMBC.2016.7590783> [doi: [10.1109/EMBC.2016.7590783](https://doi.org/10.1109/EMBC.2016.7590783)]
24. Pan G, Yan G, Qiu X, Cui J. Bleeding detection in Wireless Capsule Endoscopy based on Probabilistic Neural Network. *J Med Syst* 2011 Dec;35(6):1477-1484. [doi: [10.1007/s10916-009-9424-0](https://doi.org/10.1007/s10916-009-9424-0)] [Medline: [20703770](https://pubmed.ncbi.nlm.nih.gov/20703770/)]
25. Friberg L, Skeppholm M. Usefulness of Health Registers for detection of bleeding events in outcome studies. *Thromb Haemost* 2016 Nov 30;116(6):1131-1139. [doi: [10.1160/TH16-05-0400](https://doi.org/10.1160/TH16-05-0400)] [Medline: [27617328](https://pubmed.ncbi.nlm.nih.gov/27617328/)]

26. Gehring J, Auli M, Grangier D, Yarats D, Dauphin Y. Cornell University - arXivcs. 2017 May 8. Convolutional Sequence to Sequence Learning URL: <https://arxiv.org/abs/1705.03122> [accessed 2018-12-13] [WebCite Cache ID 74dhLoF6o]
27. Kim Y. Cornell University - arXivcs. 2014 Aug 25. Convolutional Neural Networks for Sentence Classification URL: <https://arxiv.org/abs/1408.5882> [accessed 2018-12-13] [WebCite Cache ID 74difiO2a]
28. Zhang Y, Wallace B. Cornell University - arXivcs. 2015 Oct 13. A Sensitivity Analysis of (and Practitioners? Guide to) Convolutional Neural Networks for Sentence Classification URL: <https://arxiv.org/abs/1510.03820> [accessed 2018-12-13] [WebCite Cache ID 74diiVd2O]
29. Kalchbrenner N, Grefenstette E, Blunsom P. Cornell University - arXivcs. 2014 Apr 8. A Convolutional Neural Network for Modelling Sentences URL: <https://arxiv.org/abs/1404.2188> [accessed 2018-12-13] [WebCite Cache ID 74diqirC7]
30. Cornell University - arXivcs. 2015 Dec 14. Semisupervised Autoencoder for Sentiment Analysis URL: <https://arxiv.org/abs/1512.04466> [accessed 2018-12-13] [WebCite Cache ID 74dj7usXd]
31. Cornell University - arXivcs. 2016 Jun 15. Semi-Supervised Learning for Neural Machine Translation URL: <https://arxiv.org/abs/1606.04596> [accessed 2018-12-13] [WebCite Cache ID 74djjM7s9]
32. Socher R, Huang EH, Pennington J, Ng AY, Manning CD. Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. In: NIPS'11 Proceedings of the 24th International Conference on Neural Information Processing Systems. 2011 Presented at: NIPS'11; 2011; Granada, Spain p. 801-809.
33. Gehrman S, Demnecourt F, Li Y, Carlson ET, Wu JT, Welt J, et al. Comparing deep learning and concept extraction based methods for patient phenotyping from clinical narratives. PLoS One 2018;13(2):e0192360 [FREE Full text] [doi: [10.1371/journal.pone.0192360](https://doi.org/10.1371/journal.pone.0192360)] [Medline: [29447188](https://pubmed.ncbi.nlm.nih.gov/29447188/)]
34. Chen M, Ball R, Yang L, Moradzadeh N, Chapman B, Larson D, et al. Deep Learning to Classify Radiology Free-Text Reports. Radiology 2018 Dec 13;286(3):845-852 [FREE Full text] [doi: [10.1148/radiol.2017171115](https://doi.org/10.1148/radiol.2017171115)] [Medline: [29135365](https://pubmed.ncbi.nlm.nih.gov/29135365/)]
35. Lin C, Hsu C, Lou Y, Yeh S, Lee C, Su S, et al. Artificial Intelligence Learning Semantics via External Resources for Classifying Diagnosis Codes in Discharge Notes. J Med Internet Res 2017 Nov 06;19(11):e380 [FREE Full text] [doi: [10.2196/jmir.8344](https://doi.org/10.2196/jmir.8344)] [Medline: [29109070](https://pubmed.ncbi.nlm.nih.gov/29109070/)]
36. Tran T, Kavuluru R. Predicting mental conditions based on. J Biomed Inform 2017 Nov;75S:S138-S148. [doi: [10.1016/j.jbi.2017.06.010](https://doi.org/10.1016/j.jbi.2017.06.010)] [Medline: [28606869](https://pubmed.ncbi.nlm.nih.gov/28606869/)]
37. Liu B, Dai Y, Li X, Lee W, Yu P. Building text classifiers using positive and unlabeled examples. 2003 Presented at: Proceedings of the Third IEEE International Conference on Data Mining; 2003; Melbourne, FL, USA p. 179. [doi: [10.1109/ICDM.2003.1250918](https://doi.org/10.1109/ICDM.2003.1250918)]
38. Halpern Y, Horng S, Choi Y, Sontag D. Electronic medical record phenotyping using the anchor and learn framework. J Am Med Inform Assoc 2016 Dec;23(4):731-740 [FREE Full text] [doi: [10.1093/jamia/ocw011](https://doi.org/10.1093/jamia/ocw011)] [Medline: [27107443](https://pubmed.ncbi.nlm.nih.gov/27107443/)]
39. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic differentiation in PyTorch. 2017 Oct 28 Presented at: 31st Conference on Neural Information Processing Systems; 2017; Long Beach, CA, USA.
40. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004 Jan 01;32(Database issue):D267-D270 [FREE Full text] [doi: [10.1093/nar/gkh061](https://doi.org/10.1093/nar/gkh061)] [Medline: [14681409](https://pubmed.ncbi.nlm.nih.gov/14681409/)]
41. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. JMLR. Scikit-learn: Machine Learning in Python URL: <http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html> [accessed 2018-12-17] [WebCite Cache ID 74jsCvvhho]
42. MetaMap - A Tool For Recognizing UMLS Concepts in Text. URL: <https://metamap.nlm.nih.gov/> [accessed 2018-12-17] [WebCite Cache ID 74k1GfvSp]
43. Pennington J, Socher R, Manning C. Glove: Global vectors for word representation. 2014 Presented at: Proceedings of the conference on empirical methods in natural language processing (EMNLP) .?1543; 2014; Doha, Qatar p. 1532-1543.
44. Leaman R, Khare R, Lu Z. Challenges in clinical natural language processing for automated disorder normalization. J Biomed Inform 2015 Oct;57:28-37 [FREE Full text] [doi: [10.1016/j.jbi.2015.07.010](https://doi.org/10.1016/j.jbi.2015.07.010)] [Medline: [26187250](https://pubmed.ncbi.nlm.nih.gov/26187250/)]
45. Carrell DS, Schoen RE, Leffler DA, Morris M, Rose S, Baer A, et al. Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. J Am Med Inform Assoc 2017 Sep 01;24(5):986-991. [doi: [10.1093/jamia/ocx039](https://doi.org/10.1093/jamia/ocx039)] [Medline: [28419261](https://pubmed.ncbi.nlm.nih.gov/28419261/)]

Abbreviations

- ADE:** adverse drug events
- AUC-ROC:** area under the receiver operating characteristic curve
- BiLSTM:** bidirectional long short-term memory
- BOW:** bag of words
- CNN:** convolutional neural network
- EHR:** electronic health record
- HCLA:** hybrid CNN and LSTM autoencoder
- LSTM:** long short-term memory
- NEG:** negative bleeding sample sentence.
- NLP:** natural language processing

POS: positive bleeding sample sentence.
RNN: recurrent neural network
SVM: support vector machines
UMLS: unified medical language system

Edited by G Eysenbach; submitted 15.04.18; peer-reviewed by S Gehrman, C Lin; comments to author 25.08.18; revised version received 20.10.18; accepted 01.11.18; published 08.02.19.

Please cite as:

Li R, Hu B, Liu F, Liu W, Cunningham F, McManus DD, Yu H

Detection of Bleeding Events in Electronic Health Record Notes Using Convolutional Neural Network Models Enhanced With Recurrent Neural Network Autoencoders: Deep Learning Approach

JMIR Med Inform 2019;7(1):e10788

URL: <http://medinform.jmir.org/2019/1/e10788/>

doi: [10.2196/10788](https://doi.org/10.2196/10788)

PMID: [30735140](https://pubmed.ncbi.nlm.nih.gov/30735140/)

©Rumeng Li, Baotian Hu, Feifan Liu, Weisong Liu, Francesca Cunningham, David D McManus, Hong Yu. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 08.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Use of Electronic Health Record Access and Audit Logs to Identify Physician Actions Following Noninterruptive Alert Opening: Descriptive Study

Azraa Amroze¹, BS; Terry S Field^{1,2}, DSc; Hassan Fouayzi^{1,2}, MS; Devi Sundaresan^{1,3}, MS; Laura Burns⁴, BS; Lawrence Garber^{1,3}, MD; Rajani S Sadasivam², PhD; Kathleen M Mazor^{1,2}, EdD; Jerry H Gurwitz^{1,2,3}, MD; Sarah L Cutrona^{1,2,5}, MPH, MD

¹Meyers Primary Care Institute, Worcester, MA, United States

²University of Massachusetts Medical School, Worcester, MA, United States

³Reliant Medical Group, Worcester, MA, United States

⁴University of Massachusetts Memorial Health Care, Worcester, MA, United States

⁵Edith Nourse Rogers Memorial Veterans Hospital, Veterans Health Administration, Bedford, MA, United States

Corresponding Author:

Azraa Amroze, BS

Meyers Primary Care Institute

385 Grove Street

Worcester, MA, 01605

United States

Phone: 1 774 275 6030

Fax: 1 508 453 8193

Email: Azraa.Amroze@meyersprimary.org

Abstract

Background: Electronic health record (EHR) access and audit logs record behaviors of providers as they navigate the EHR. These data can be used to better understand provider responses to EHR-based clinical decision support (CDS), shedding light on whether and why CDS is effective.

Objective: This study aimed to determine the feasibility of using EHR access and audit logs to track primary care physicians' (PCPs') opening of and response to noninterruptive alerts delivered to EHR InBaskets.

Methods: We conducted a descriptive study to assess the use of EHR log data to track provider behavior. We analyzed data recorded following opening of 799 noninterruptive alerts sent to 75 PCPs' InBaskets through a prior randomized controlled trial. Three types of alerts highlighted new medication concerns for older patients' posthospital discharge: information only (n=593), medication recommendations (n=37), and test recommendations (n=169). We sought log data to identify the person opening the alert and the timing and type of PCPs' follow-up EHR actions (immediate vs by the end of the following day). We performed multivariate analyses examining associations between alert type, patient characteristics, provider characteristics, and contextual factors and likelihood of immediate or subsequent PCP action (general, medication-specific, or laboratory-specific actions). We describe challenges and strategies for log data use.

Results: We successfully identified the required data in EHR access and audit logs. More than three-quarters of alerts (78.5%, 627/799) were opened by the PCP to whom they were directed, allowing us to assess immediate PCP action; of these, 208 alerts were followed by immediate action. Expanding on our analyses to include alerts opened by staff or covering physicians, we found that an additional 330 of the 799 alerts demonstrated PCP action by the end of the following day. The remaining 261 alerts showed no PCP action. Compared to information-only alerts, the odds ratio (OR) of immediate action was 4.03 (95% CI 1.67-9.72) for medication-recommendation and 2.14 (95% CI 1.38-3.32) for test-recommendation alerts. Compared to information-only alerts, ORs of medication-specific action by end of the following day were significantly greater for medication recommendations (5.59; 95% CI 2.42-12.94) and test recommendations (1.71; 95% CI 1.09-2.68). We found a similar pattern for OR of laboratory-specific action. We encountered 2 main challenges: (1) Capturing a historical snapshot of EHR status (number of InBasket messages at time of alert delivery) required incorporation of data generated many months prior with longitudinal follow-up. (2) Accurately

interpreting data elements required iterative work by a physician/data manager team taking action within the EHR and then examining audit logs to identify corresponding documentation.

Conclusions: EHR log data could inform future efforts and provide valuable information during development and refinement of CDS interventions. To address challenges, use of these data should be planned before implementing an EHR-based study.

(*JMIR Med Inform* 2019;7(1):e12650) doi:[10.2196/12650](https://doi.org/10.2196/12650)

KEYWORDS

electronic health records; health services research; health information technology; health care communication

Introduction

Audit and access logs in the electronic health record (EHR) have primarily been used for security purposes [1-3], but recent studies [4-7] indicate that a broad range of additional research and clinical questions may be answered using this relatively untapped data source. Access logs contain time-stamped recordings of who accessed the EHR and what part was accessed [5]. Audit logs record chronological activity in the EHR, tracking actions such as what data were created, reviewed, or changed by the user [3]. These log data, though potentially complex to retrieve and interpret, can be used by researchers and care teams to identify targets for clinical interventions as well as care quality and safety-improvement efforts [4,8].

Data from access and audit logs can be used to better understand how various forms of clinical decision support (CDS) impact physician behavior. This is especially relevant for noninterruptive EHR alerts, where assessments of the effectiveness of these alerts have found mixed results [9-12]. In contrast to alerts that “pop-up” and interrupt workflow when a specific EHR action is taken, noninterruptive alerts deliver information to an EHR InBasket, requiring the receiver to open and review them. This information may include warnings about out-of-range test results [9,13,14], abnormal findings on diagnostic imaging and screening [10,15-17], important changes in patient health [18], and safety concerns during transition of care [19]. However, these alerts must compete for the attention of physicians who are often overwhelmed by the number of electronic notifications they receive, read, and respond to while managing direct patient care [4,20-23].

Access logs can provide valuable insight into the precise time of alert opening, even if the alert is opened many days postdelivery; logs can also help pinpoint actions following opening that might otherwise be difficult to link back to a particular CDS intervention. Although causality of EHR actions can be difficult to prove, there is value to demonstrating close temporal proximity of an EHR action after alert opening. Such knowledge can shed light on the mechanisms by which an EHR-based intervention is (or is not) effective.

The effectiveness of a noninterruptive alerting system was tested by our team in a previous randomized trial (trial registration: ClinicalTrials.gov NCT00611091) [19]. Locally developed alerts regarding older adults' posthospital discharge were sent to their primary care physicians (PCPs) through the EHR [24,25]. These alerts highlighted what our team believed to be actionable medication safety concerns. The trial assessed whether these alerts increased outpatient visits and reduced

rehospitalization [19]. The alerts did not improve either of these measures. To better understand this outcome, we conducted this descriptive study using audit and access log data to track actions taken by physicians in the EHR following alert delivery. In this paper, we present findings from this analysis of physician action. We also describe the availability and level of detail of information in these logs and review challenges and strategies for success.

Methods

Design

This study was based on data from the intervention arm of a prior randomized controlled trial, as mentioned above [19]. In that trial, an EHR-based intervention was implemented using locally developed noninterruptive alerts to highlight postdischarge medication safety concerns for older patients. The inpatient facility to which the health care system admitted its patients used a different EHR than that used by the health care system. An interface engine was linked to the hospital's admission, discharge, and transfer registration system. The hospital transmitted information including discharge dates for the health care system's patients; these data were automatically incorporated into the health care system's EHR. Informed by health plan data reflecting new prescriptions filled, alerts were then generated for patients. These alerts were designed to convey actionable time-sensitive concerns pertaining to high-risk drugs (known to result in more adverse drug events) and new drugs. Over the 4 months prior to implementation, two physicians from the health care system met weekly to review every alert generated, suggesting modifications to ensure that alerts would be perceived as necessary, useful, and brief. Alerts that they considered inactionable were eliminated. At the time of hospital discharge, patients were randomized to the intervention (EHR alerts sent to the patient's PCP if medication concerns were identified) or usual follow-up care. The primary goal of the intervention was to decrease rehospitalization rates.

Setting

This study was conducted at a large Massachusetts multispecialty medical group with 217 physicians at 15 sites throughout Central Massachusetts. All sites used an EHR from Epic Systems Corporation.

Study Sample

When the original trial was conducted in 2011, approximately 140,000 adults received primary care through this medical group, of which approximately 24,000 were aged ≥ 65 years and were members of a local health plan. Patients who were

members of this local health plan with primary care providers within this medical group and were discharged to home from the primary hospital used by the medical group from August 26, 2010, to August 25, 2011, were included in the trial. A total of 1282 patients were randomized to the intervention arm. In our analysis, we included only the 799 patient discharges for which PCP medication alerts were triggered (corresponding to 713 patients). These alerts were sent to 75 PCPs.

The Institutional Review Boards at Reliant Medical Group and the University of Massachusetts Medical School approved this study, and waiver of consent was obtained.

Variables and Data Sources

Data available from the original trial included details about the timing and content of alerts, the scrambled identifier of the physician to whom the alerts were sent, and additional PCP characteristics including age (<50 vs ≥ 50 years), gender, and specialty (internal medicine, family medicine, PCP without a Doctor of Medicine degree such as a nurse practitioner, and subspecialist acting as PCP). To approximate full- versus part-time status of the physicians, we categorized the total number of patient encounters for each physician during the year prior to alert delivery in quartiles (0-2326, 2327-2783, 2784-3173, and >3173) across the 75 PCPs to whom alerts were sent. We collected the following information about the relevant patients: age (65-74, 75-84, and ≥ 85 years), gender, Charlson comorbidity index score (categorized as 0, 1, 2, and ≥ 3), dates and length (≤ 2 days, 3 days, and ≥ 4 days) of the related hospitalization, number of office visits in the past year (≤ 6 , 7-11, 12-18, and >18), patient outcomes, and scrambled patient identifiers. Our research team grouped all the alerts sent during the intervention into three categories: information about new and high-risk medications at the time of hospital discharge ("information only"), recommendation to cancel or modify doses of medications ("medication recommendation"), and recommendation to order tests—laboratory tests or, in a few cases, eye exams—to monitor the impact of high-risk medications or titrate their doses ("test recommendation").

For this descriptive study, we sought a range of additional data elements from the EHR access and audit logs (described below) to provide information on the opening of alerts and track specific physician actions within the EHR. We captured the timing of alert opening relative to alert delivery (≤ 24 hours, 24-48 hours, and >48 hours) within office hours (8 AM-5 PM Monday through Friday) as compared to outside office hours. We also identified the user opening the alert (the intended PCP vs a staff member or provider other than the intended PCP).

We gathered data on the following variables pertaining to the PCP's InBasket load categorized by quartile: total number of notifications in the InBasket at the time of alert delivery (≤ 42 , 43-69, 70-157, and >157), number of unopened notifications in the InBasket at the time of alert delivery (0, 1-4, 5-9, and >9), and number of notifications delivered to the InBasket in the 7 days prior to alert delivery (≤ 344 , 345-453, 454-546, and >546). Finally, we compared alerts delivered on Saturdays to those delivered on all other days, because those arriving on Saturday were the only alerts for which the 24 hours postdelivery did not include any weekday time.

Our goal was to identify factors associated with the alerts that impacted physician review of patient information (eg, review of EHR information related to prescribed medications, laboratory test results, and laboratory orders) or physician orders for medications or laboratory tests.

Identifying Primary Care Physicians' Actions in the Electronic Health Record Through Use of Access and Audit Logs

We analyzed EHR log data corresponding to the 2-day period following opening of 799 alerts (593 information only, 37 medication recommendations, and 169 test recommendations) sent to 75 PCPs' InBaskets. These three types of alerts highlighted new medication concerns for older patients posthospital discharge; all were intended to prompt review of the recently hospitalized patient's chart within the EHR by the PCP. Expected actions for test recommendations included viewing test results or ordering tests. Expected actions for medication recommendations included viewing medication lists and discontinuing or ordering medications. For information-only alerts, appropriate action was left to the judgment of the PCP, but expected actions were similar to those for medication recommendations and included performance of medication reconciliation, which would entail viewing medication lists and possibly discontinuing or ordering medications. In all cases, reviews of other parts of the patient chart could be expected in support of information gathering and decision making.

In order to retrieve data on the 2-day period following alert opening, the data manager needed to first identify the location and format of data generated by the locally developed alert system. This work was accomplished through careful coordination between a physician with EHR access and a data manager. Simulated patients were created, and alerts were triggered for these simulated patients. Following this, a collaborating physician from the medical group opened the alerts and then opened and viewed various sections within the EHR corresponding to the simulated patient's chart. The data manager used the collaborating physician and simulated patient identifiers together with the dates and times of triggering activities to locate these actions within the EHR's logs and tables. This allowed her to identify the code generated by the locally developed alert system, indicating that an alert had been sent. From that base, the unique identifiers generated for each delivered alert could be extracted. Using the audit log table, we then obtained time stamps for each alert opened (hour, minute, and second) and an identifier for the person opening the alert in order to track actions. Alerts could be opened by the PCP to whom they were addressed or other members of the care team.

Immediate Actions

We identified the first time the alert was opened by the addressed PCP. The InBasket view of the EHR version used during the study period only displayed the alert message and several clickable buttons that served as direct links to summaries of sections of the corresponding patient's chart; thus, navigation to other sections of the EHR was required to gain any additional information. After opening the alert, some PCPs navigated to sections of the EHR for the patient triggering the alert, whereas some moved on to a different alert or the electronic record for

a different patient. Using time stamps, we captured the time spent by the PCP on the alert itself, and we captured the total time spent on the relevant patient. Total time included the duration during which the alert was displayed in the EHR plus the time the PCP spent navigating sections of the EHR corresponding to the relevant patient.

Using the EHR access log table, we identified each action taken by the provider during the 5-minute period immediately following alert opening. Data elements from each action provided information on the exact time (hour, minute, seconds) of the action as well as codes indicating the specific section of the EHR opened and the patient to whom it related. Codes indicating the component of the EHR accessed by the PCP were complex, requiring additional physician-data manager collaboration to categorize. Actions that we considered as viewing relevant patient information included opening a section of the electronic medical record (medications, laboratory, orders, results, encounters, demographic information, other clinical information, nonclinical information, and information entry) or choosing one of several options on the alert that served as direct links to summaries of components of that patient's record. Actions that we considered as not viewing relevant patient information included opening a notification related to a different patient, opening a section of a different patient's medical record, or doing nothing further in the EHR for 5 minutes.

Subsequent Relevant Actions

Considering that PCPs might be opening alerts briefly in between patient visits and might return to the patient record later to address issues raised in the alert, we sought to broaden the window for tracking PCP's actions. For our analysis of subsequent relevant actions taken by the PCP, the time window studied included the day of alert opening and the following day. Extracted data included a timestamp as well as codes indicating the EHR component accessed (categorized in the same manner as with immediate action analyses described above). In addition to the information extracted from the access log table, we captured evidence of orders placed for medications (new medication, change, or discontinuation) and laboratory tests. We focused only on actions taken in the EHR for the patient of interest (eg, the patient for whom the alert was triggered).

We categorized PCP actions under the broad heading of general action (includes viewing any patient data, creating documentation, and placing any orders). Within this broad group, we defined subcategories: medication-specific action (viewing the patient's medication list or ordering a medication) and laboratory-specific action (viewing or ordering laboratory tests).

Contextual Factors

We also examined contextual factors potentially relevant to alert opening and subsequent EHR action. These included the number of messages and unopened messages in the physician's InBasket as well as the flow of messages during the prior week. Capturing a snapshot of the EHR status at a historical point in time (eg, number of InBasket messages at the time of alert delivery) presented challenges. To recreate the InBasket at the moment of alert delivery, we began by assembling all the

messages delivered to a specific physician's InBasket during the 1-year period prior to alert delivery. We then eliminated any messages for which the record indicated that the message was "completed" (eg, deleted from the InBasket) prior to the date of alert delivery. This provided a count of the number of messages remaining in the InBasket. To calculate the number of messages that remained unopened at the time of alert delivery, we identified those for which there was no log entry indicating opening on or prior to the alert delivery date. We also calculated the total number of notifications that arrived in the InBasket during the week prior to alert delivery.

Data Analysis

Immediate Actions

For assessment of the immediate response during the 5 minutes after alert opening, we focused only on those cases in which the physician to whom the alert was sent was the first person to open the alert. We categorized the immediate next action by the physician as one related or not related to the relevant patient. We used bivariate analyses to assess the relationship between performing an immediate relevant action and alert type, timing of alert opening relative to alert delivery, provider and patient characteristics, and contextual factors. We performed multivariate analyses to obtain odds ratios (OR) in the presence of more than one variable. Since some PCPs received multiple alerts over the 1-year period, we used generalized estimating equations with a logit link and a binomial distribution to account for clustering of measures within PCPs. To estimate the total time attributable to the alert in the 5 minutes following opening, we calculated the length of time spent by the PCP with the alert open (ie, displayed on the EHR computer screen) and combined this with the total time (immediately following alert opening) spent viewing sections of the EHR for that patient.

Subsequent Relevant Actions

When assessing factors related to relevant subsequent actions, we included all first openings of each alert irrespective of whether the addressed physician or a covering physician or staff member opened the alert, but tracked behavior of the addressed physician. Similar to the process for the immediate actions, we performed bivariate and multivariate analyses using generalized estimating equation models to assess the relationship between performing subsequent relevant actions and the covariates mentioned above. We also assessed the effect of alert opening by the addressed physician on the likelihood of subsequent actions.

Results

Overview

In the overall group of 799 alerts opened, 627 (78.5%) were opened by the addressed PCP. Of this subgroup, we were able to track actions immediately following PCP alert opening for 616 alerts (11 alerts had no available data). The analysis was then expanded to include the remaining 172 (21.5%) alerts opened by staff or a provider other than the PCP to track actions over the day of alert opening and the following day. Alerts that did not show evidence of action within this timeframe were

classified as alerts not followed by timely PCP action (Figure 1).

Immediate Actions

Among the 616 tracked alerts, 208 (33.8%) were immediately followed by viewing of the relevant patient’s EHR (Table 1). Of 445 information-only alerts, 28.1% (125/445) were followed by immediate viewing of the patient’s EHR, as were 54.8% (17/31) of the medication-recommendation alerts and 47.1% (66/140) of the test-recommendation alerts.

In multivariate analyses predicting immediate viewing of the patient’s electronic information, the only factor that reached

statistical significance was the type of alert. As compared to information-only alerts (reference group), the odds ratio (OR) of immediate action for medication-recommendation alerts was 4.03 (95% CI 1.67-9.72); the OR for test-recommendation alerts was 2.14 (95% CI 1.38-3.32). In the 5 minutes postopening, the mean time PCPs spent on the patient of interest—with the alert on display and while navigating relevant patient’s EHR—was 106 seconds (median, 64 seconds). Alerts not immediately followed by viewing of the relevant patient’s EHR were kept on display for a mean time of 26 seconds (median, 15 seconds; Table 2).

Figure 1. Overview of alert opening. Asterisk indicates that 11 alerts had no data available on immediate action. PCP: primary care physician.

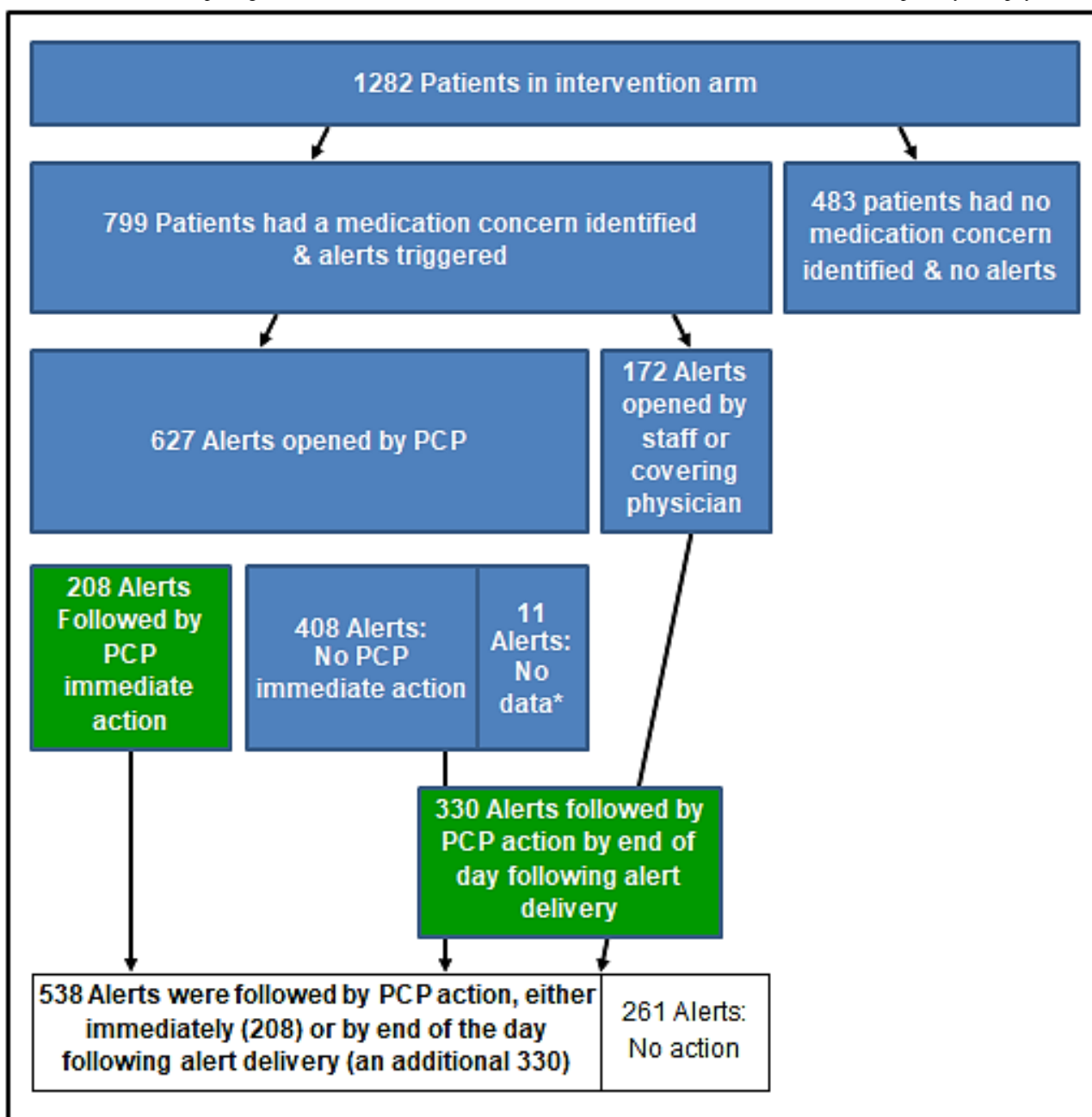


Table 1. Immediate electronic health record actions taken by primary care providers after opening noninterruptive alerts. We identified the first time the alert was opened by the addressed primary care physician and extracted from the electronic health record access log table data on the first action taken by the provider during the 5-minute period immediately following alert opening.

Characteristics of alerts, physicians, patients, and contextual factors	N	Immediate action ^a taken in the record of the patient who triggered the alert, n (%)	No immediate action ^b taken in the record of the patient who the triggered alert, n (%)
Total	616	208 (33.8)	408 (66.2)
Alert type			
Information only	445	125 (60.1)	320 (78.4)
Medication recommendation	31	17 (8.2)	14 (3.4)
Test recommendation	140	66 (31.7)	74 (18.1)
Time to alert opening			
Opened ≤1 hour	119	32 (15.4)	87 (21.3)
Opened ≤24 hours	252	88 (42.3)	164 (40.2)
Opened >24 hours	245	88 (42.3)	157 (38.5)
Characteristics of the primary care physician			
Gender			
Female	204	72 (34.6)	132 (32.4)
Male	412	136 (65.4)	276 (67.6)
Number of patient encounters in the study year (quartiles)^c			
>0 and ≤2326	42	19 (9.1)	23 (5.6)
>2326 and ≤2783	115	33 (15.9)	82 (20.1)
>2783 and ≤3173	203	70 (33.7)	133 (32.6)
>3173	256	86 (41.3)	170 (41.7)
Patient characteristic - Charlson comorbidity score			
0	58	18 (8.7)	40 (9.8)
1	67	20 (9.6)	47 (11.5)
≥2	491	170 (81.7)	321 (78.7)
Contextual factors			
Opening within office hours			
No: Opened outside office hours	312	110 (52.9)	202 (49.5)
Yes: Opened 8 AM to 5 PM Mon-Fri	304	98 (47.1)	206 (50.5)
Total number of notifications (opened + unopened) in InBasket at time of alert delivery			
≤42	159	42 (20.2)	117 (28.7)
>42 and ≤69	152	50 (24)	102 (25)
>69 and ≤157	153	62 (29.8)	91 (22.3)
>157	152	54 (26)	98 (24)
Number of unopened notifications in InBasket at the time of alert delivery			
≤0	189	60 (28.8)	129 (31.6)
>0 and ≤4	135	43 (20.7)	92 (22.5)
>4 and ≤9	145	50 (24)	95 (23.3)
>9	147	55 (26.4)	92 (22.5)
Notification count in the week prior to alert delivery			
≤344	146	54 (26)	92 (22.5)
>344 and ≤453	159	48 (23.1)	111 (27.2)
>453 and ≤546	157	42 (20.2)	115 (28.2)

Characteristics of alerts, physicians, patients, and contextual factors	N	Immediate action ^a taken in the record of the patient who triggered the alert, n (%)	No immediate action ^b taken in the record of the patient who triggered the alert, n (%)
>546	154	64 (30.8)	90 (22.1)

^aActions that we considered as viewing relevant patient information included opening a section of the electronic medical record (medications, laboratory, orders, results, encounters, demographic information, other clinical information, nonclinical information, and information entry) or choosing one of several options on the alert that served as direct links to summaries of components of that patient's record.

^bActions that we considered as not viewing relevant patient information included opening a notification related to a different patient, opening a section of a different patient's medical record, or doing nothing further in the electronic health record for 5 minutes.

^cWe categorized the total number of patient encounters for each physician during the year prior to alert delivery in quartiles across the 75 primary care physicians to whom alerts were sent to approximate full- vs part-time status.

Table 2. Time spent by the primary care physician in electronic charts of patients triggering an alert during the 5 minutes postalert opening.

Time spent in the electronic chart, by category	Time (seconds)	
	Mean	Median
Time spent in the chart of the relevant patient (n=616 alerts)	55	22
Time spent in the chart when the primary care physician took immediate action in the record of the patient who triggered the alert ^a (n=208)	106	64
Time spent in the chart when the primary care physician took no immediate action in the record of the patient who triggered the alert ^b (n=408)	26	15
Time spent in the chart of the relevant patient, by alert type		
Information only (n=445)	42	17
Medication recommendation (n=31)	81	51
Test recommendation (n=140)	90	42

^aActions that we considered as viewing relevant patient information included opening a section of the electronic medical record (medications, laboratory, orders, results, encounters, demographic information, other clinical information, nonclinical information, and information entry) or choosing one of several options on the alert that served as direct links to summaries of components of that patient's record.

^bActions that we considered as not viewing relevant patient information included opening a notification related to a different patient, opening a section of a different patient's medical record, or doing nothing further in the EHR for 5 minutes. Thus, this number represents an estimate of time spent viewing an alert.

Subsequent Relevant Actions

In the overall group of 799 alerts opened, 538 (67.3%) were followed by PCP-related actions during the day of opening or the next day (Table 3). Among information-only alerts, 64.4% (382/593) were followed by a general action in the EHR; 16.0% (95/593), by a medication-specific action; and 16.9% (100/593), by a laboratory-specific action. Among medication-recommendation alerts, 78.4% (29/37) were followed by a general action; 48.6% (18/37), by a medication-specific action; and 54.1% (20/37), by a laboratory-specific action. Among test-recommendation alerts, 75.1% (127/169) were followed by a general action, 24.3% (41/169) were followed by a medication-specific action, and 35.5% (60/169) were followed by a laboratory-specific action.

On multivariate analysis (Table 3), alerts containing specific instructions for the PCP (medication recommendations or test recommendations) were significantly more likely to be

associated with subsequent action by the PCP compared to information-only alerts. Compared to information-only alerts, the ORs of medication-specific action were significantly greater for medication recommendations (OR: 5.59; 95% CI 2.42-12.94) and test recommendations (OR: 1.71; 95% CI 1.09-2.68). Likewise, the ORs of laboratory-specific action were significantly greater for medication recommendations (OR: 7.37; 95% CI 3.64-14.97) and test recommendations (OR: 2.75; 95% CI 1.73-4.38).

There was no significant association between the timing of alert opening and responsive action. As expected, subsequent relevant action by the PCP was more likely for alerts opened by the addressed PCP as compared to alerts opened by staff or another provider, although only the association with medication-specific action reached statistical significance (general action, OR: 1.41; 95% CI 0.81-2.45; medication-specific action, OR: 2.13; 95% CI 1.35-3.37; laboratory-specific action, OR: 1.65; 95% CI 0.99-2.73).

Table 3. Multivariate analysis results for factors associated with primary care providers' subsequent action following opening of noninterruptive alerts. Alerts were opened by the addressed primary care provider, staff, or provider other than the addressed primary care provider. The primary care provider's actions in the electronic health record related to the relevant patient were tracked over the day of alert opening and the following day.

Characteristics of alerts, physicians, patients, and contextual factors	Total alerts N (%)	General action ^a in electronic health record taken by PCP ^b		Medication-specific action ^c taken by PCP		Laboratory-specific action ^d taken by PCP	
		Action is taken, N (%)	AOR ^e (95% CI)	Action is taken, N (%)	AOR (95% CI)	Action is taken, N (%)	AOR (95% CI)
Total	799	538 (67.3)	N/A ^f	154 (19.3)	N/A	180 (22.5)	N/A
Alert type							
Information only	593 (74.2)	382 (71)	Reference	95 (61.7)	Reference	100 (55.6)	Reference
Medication recommendation ^g	37 (4.6)	29 (5.4)	2.0 (0.9- 4.8)	18 (11.7)	5.6 (2.4- 12.9)	20 (11.1)	7.4 (3.6- 15.0)
Test recommendation ^h	169 (21.2)	127 (23.6)	1.7 (1.1- 2.5)	41 (26.6)	1.7 (1.1- 2.7)	60 (33.3)	2.8 (1.7- 4.4)
Time to alert opening							
Opened ≤24 h after delivery	472 (59.1)	308 (57.2)	Reference	88 (57.1)	Reference	101 (56.1)	Reference
Opened >24 h and ≤48 h after delivery	137 (17.1)	97 (18)	1.2 (0.6- 2.4)	27 (17.5)	1.2 (0.7- 1.9)	28 (15.6)	1.1 (0.6- 2.0)
Opened >48 h after delivery	190 (23.8)	133 (24.7)	1.2 (0.7- 2.2)	39 (25.3)	1.1 (0.6- 1.8)	51 (28.3)	1.4 (0.8- 2.4)
Opened by PCP/other staff							
Opened by staff/provider other than PCP	172 (21.5)	100 (18.6)	Reference	17 (11)	Reference	23 (12.8)	Reference
Opened by PCP	627 (78.5)	438 (81.4)	1.4 (0.8-2.5)	137 (89)	2.1 (1.4-3.4)	157 (87.2)	1.7 (1.0-2.7)
PCP characteristics							
Age							
<50 years	296 (37.0)	203 (37.7)	Reference	63 (40.9)	Reference	80 (44.4)	Reference
≥50 years	503 (63.0)	335 (62.3)	1.0 (0.7- 1.4)	91 (59.1)	0.6 (0.4- 1.1)	100 (55.6)	0.6 (0.4- 0.9)
Gender							
Female	264 (33.0)	182 (33.8)	Reference	54 (35.1)	Reference	71 (39.4)	Reference
Male	535 (67.0)	356 (66.2)	0.6 (0.4- 1.0)	100 (64.9)	0.7 (0.4- 1.3)	109 (60.6)	0.5 (0.3- 0.7)
Number of patient encounters in study year (quartiles)ⁱ							
>0 and ≤2326	80 (10.0)	58 (10.8)	Reference	8 (5.2)	Reference	14 (7.8)	Reference
>2326 and ≤2783	163 (20.4)	94 (17.5)	0.6 (0.3- 1.4)	18 (11.7)	0.5 (0.4- 2.8)	30 (16.7)	1.0 (0.4- 2.2)
>2783 and ≤3173	247 (30.9)	168 (31.2)	1.0 (0.5- 2.1)	54 (35.1)	1.5 (1.2- 8.0)	59 (32.8)	1.9 (0.9- 4.0)
>3173	309 (38.7)	218 (40.5)	1.5 (0.7- 3.3)	74 (48.1)	2.5 (1.5- 13.3)	77 (42.8)	2.7 (1.3- 5.8)
Specialty							
Internal medicine	118 (14.8)	445 (82.7)	Reference	129 (83.8)	Reference	155 (86.1)	Reference
Family medicine	661 (82.7)	79 (14.7)	0.7 (0.4- 1.2)	22 (14.3)	0.9 (0.4- 1.8)	20 (11.1)	0.6 (0.4- 0.9)
Non-MD ^j PCP	11 (1.4)	6 (1.1)	0.3 (0.1- 1.2)	2 (1.3)	0.6 (0.2- 1.9)	3 (1.7)	0.6 (0.3- 1.4)
Subspecialty	9 (1.1)	8 (1.5)	3.0 (1.0- 9.6)	1 (0.6)	1.4 (0.4- 5.3)	2 (1.1)	2.1 (0.7- 6.0)
Patient characteristics							
Age							
65-74 years	255 (31.9)	171 (31.8)	Reference	55 (35.7)	Reference	56 (31.1)	Reference
75-84 years	349 (43.7)	241 (44.8)	1.1 (0.7- 1.6)	57 (37)	0.8 (0.5- 1.1)	77 (42.8)	1.0 (0.6- 1.4)
≥85 years	195 (24.4)	126 (23.4)	0.9 (0.5- 1.4)	42 (27.3)	1.0 (0.6- 1.5)	47 (26.1)	1.0 (0.6- 1.6)
Gender							

Characteristics of alerts, physicians, patients, and contextual factors	Total alerts	General action ^a in electronic health record taken by PCP ^b		Medication-specific action ^c taken by PCP		Laboratory-specific action ^d taken by PCP	
	N (%)	Action is taken, N (%)	AOR ^e (95% CI)	Action is taken, N (%)	AOR (95% CI)	Action is taken, N (%)	AOR (95% CI)
Female	418 (52.3)	284 (52.8)	Reference	81 (52.6)	Reference	93 (51.7)	Reference
Male	381 (47.7)	254 (47.2)	0.9 (0.7- 1.3)	73 (47.4)	1.1 (0.8- 1.6)	87 (48.3)	1.3 (0.9- 1.8)
Number of office visits in the previous 12 months							
≤6 visits	205 (25.7)	125 (23.2)	Reference	41 (26.6)	Reference	39 (21.7)	Reference
>6 and ≤11 visits	219 (27.4)	142 (26.4)	1.3 (0.9- 2.0)	36 (23.4)	0.9 (0.5- 1.6)	44 (24.4)	1.0 (0.6- 1.7)
>11 and ≤18 visits	194 (24.3)	144 (26.8)	2.1 (1.3- 3.6)	42 (27.3)	1.4 (0.8- 2.3)	48 (26.7)	1.4 (0.6- 1.7)
>18 visits	181 (22.7)	127 (23.6)	1.7 (1.0- 2.7)	35 (22.7)	1.2 (0.6- 2.1)	49 (27.2)	1.5 (0.9- 2.5)
Charlson comorbidity score							
0	78 (9.8)	51 (9.5)	Reference	17 (11)	Reference	11 (6.1)	Reference
1	92 (11.5)	56 (10.4)	0.7 (0.3- 1.4)	16 (10.4)	0.7 (0.3- 1.7)	17 (9.4)	1.3 (0.5- 3.1)
2	118 (14.8)	80 (14.9)	0.9 (0.4- 1.8)	21 (13.6)	0.7 (0.3- 1.6)	25 (13.9)	1.6 (0.6- 3.9)
≥3	511 (64.0)	351 (65.2)	0.8 (0.4- 1.4)	100 (64.9)	0.6 (0.3- 1.3)	127 (70.6)	1.3 (0.6- 2.8)
Length of stay							
≤2 days	360 (45.1)	229 (42.6)	Reference	65 (42.2)	Reference	74 (41.1)	Reference
3 days	281 (35.2)	199 (37)	1.4 (1.0- 2.1)	57 (37)	1.1 (0.8- 1.6)	67 (37.2)	1.2 (0.9- 1.7)
≥4 days	158 (19.8)	110 (20.4)	1.5 (0.9- 2.5)	32 (20.8)	1.3 (0.8- 2.1)	39 (21.7)	1.5 (0.9- 2.4)
Contextual factors							
Total number of notifications in InBasket at the time of alert delivery							
≤42	207 (25.9)	146 (27.1)	Reference	43 (27.9)	Reference	51 (28.3)	Reference
>42 and ≤69	194 (24.3)	134 (24.9)	0.9 (0.5- 1.5)	32 (20.8)	0.9 (0.5- 1.5)	34 (18.9)	0.7 (0.5- 1.1)
>69 and ≤157	199 (24.9)	135 (25.1)	0.7 (0.4- 1.2)	37 (24)	1.0 (0.6- 1.8)	51 (28.3)	1.4 (0.9- 2.0)
>157	199 (24.9)	123 (22.9)	0.5 (0.3-1.0)	42 (27.3)	1.0 (0.7- 2.4)	44 (24.4)	0.9 (0.6- 1.5)
Number of unopened notifications in InBasket at the time of alert delivery							
≤0	251 (31.4)	165 (30.7)	Reference	50 (32.5)	Reference	57 (31.7)	Reference
>0 and ≤4	183 (22.9)	122 (22.7)	1.0 (0.7- 1.5)	34 (22.1)	0.8 (0.5-1.3)	42 (23.3)	0.9 (0.6- 1.5)
>4 and ≤9	185 (23.2)	127 (23.6)	1.2 (0.7- 2.0)	32 (20.8)	0.7 (0.5- 1.0)	45 (25)	0.8 (0.5- 1.4)
>9	180 (22.5)	124 (23)	1.5 (0.9- 2.5)	38 (24.7)	1.2 (0.8- 1.8)	36 (20)	0.8 (0.5- 1.2)
Notification count in the week prior to alert delivery							
≤344	200 (25.0)	146 (27.1)	Reference	39 (25.3)	Reference	55 (30.6)	Reference
>344 and ≤453	201 (25.2)	127 (23.6)	0.7 (0.4- 1.0)	43 (27.9)	0.9 (0.6- 1.4)	49 (27.2)	0.6 (0.4- 1.1)
>453 and ≤546	199 (24.9)	135 (25.1)	0.7 (0.5- 1.1)	35 (22.7)	0.6 (0.4- 0.9)	35 (19.4)	0.4 (0.3- 0.8)
>546	199 (24.9)	130 (24.2)	0.7 (0.4- 1.1)	37 (24)	0.6 (0.4- 1.0)	41 (22.8)	0.6 (0.4- 1.0)
Day of the week alert sent							
All other days	633 (79.2)	418 (77.7)	Reference	119 (77.3)	Reference	143 (79.4)	Reference

Characteristics of alerts, physicians, patients, and contextual factors	Total alerts	General action ^a in electronic health record taken by PCP ^b		Medication-specific action ^c taken by PCP		Laboratory-specific action ^d taken by PCP	
	N (%)	Action is taken, N (%)	AOR ^e (95% CI)	Action is taken, N (%)	AOR (95% CI)	Action is taken, N (%)	AOR (95% CI)
Saturday	166 (20.8)	120 (22.3)	1.3 (0.7- 2.2)	35 (22.7)	1.2 (0.7- 1.9)	37 (20.6)	0.9 (0.6- 1.6)

^aGeneral Action in electronic health record includes opening a section of the electronic medical record (medications, laboratory, orders, results, encounters, demographic information, other clinical information, nonclinical information, and information entry) or choosing one of several options on the alert that served as direct links to summaries of components of that patient's record.

^bMedication-specific action includes medication list viewing and ordering.

^cLaboratory-specific action includes laboratory viewing and ordering.

^dPCP: primary care physician.

^eAOR: adjusted odds ratio.

^fN/A: not applicable.

^gMedication recommendations were automated electronic health record InBasket alerts that contained warnings about interactions or recommendations for dose changes.

^hTest recommendations were automated electronic health record InBasket alerts identifying the need for laboratory monitoring for high-risk medications.

ⁱWe categorized the total number of patient encounters for each physician during the year prior to alert delivery in quartiles across the 75 primary care physicians to whom alerts were sent to approximate full- vs part-time status.

^jMD: doctor of medicine.

Contextual factors showed no consistent patterns. There was some suggestion of a stronger likelihood for responsive action if the relevant patient had more than 11 office visits in the past year. Responsive actions were also more likely taken by physicians with more patient encounters and less likely by physicians with total number of InBasket notifications at the time of alert delivery in the top quartile (>157 alerts).

Electronic Health Record Access and Audit Logs

Challenges

We were able to identify all the data that were required to track provider behavior in the EHR access and audit logs; however, a substantial effort was required to identify and interpret these data. We encountered two main challenges in data retrieval and interpretation. As detailed below, these challenges pertained to capturing a historical snapshot (eg, number of InBasket messages at a given time in the past) and interpreting data elements from access and audit logs in the absence of clear data documentation.

Capturing a Historical Snapshot of Electronic Health Record Status

Our access logs do not keep successive records of evolving EHR elements (examples include InBasket message lists or medication lists, both of which have elements added and deleted over time). Thus, in order to capture the status of a provider's InBasket retrospectively, we had to retrieve data generated many months prior and couple this with longitudinal follow-up data to determine the status at the time of interest. Specifically, to recreate a list of InBasket messages present at the time of alert delivery, we obtained an estimate of the date on which the earliest arriving messages were generated (we went back 1 year) and then captured all the messages generated from that point forward. We honed this large list using status data reflecting whether the message was opened and whether it had been "completed" (eg, deleted) prior to the date of alert delivery. Messages that were marked as completed (and those postponed

to return at a future date) were removed; the remaining were concluded to be messages present in the InBasket at the time of alert delivery.

Interpretation of Data Elements Required an Iterative Approach With Physician-Data Manager Collaboration

We encountered challenges in deciphering the variable names for data elements drawn from the audit logs. These data had not previously been used for research purposes, and therefore, we lacked clear documentation for the interpretation of many variables. Accurately interpreting data elements required iterative work by the physician/data manager team. Using a simulated patient record in the EHR, our physician collaborator performed actions through the physician user interface and then worked with the data manager to compare results captured simultaneously through the audit log. Both a physician's perspective (providing insight into routine actions taken by physicians within an EHR) and the data manager's informatics knowledge (and access to data tables) were essential for this effort.

Discussion

Our descriptive study assessed the use of EHR log data to track provider behavior following opening of noninterruptive medication alerts. More than three-quarters of alerts (78.5%; 627/799) were opened by the PCP to whom they were directed, allowing us to assess immediate PCP action; of these, 208 alerts were followed by immediate action. Expanding our analysis to include alerts opened by staff or covering physicians, we found that an additional 330/799 alerts showed evidence of action by the end of the following day. The remaining 261 alerts showed no evidence of PCP action. Compared to information-only alerts, alerts containing specific instructions (including medication and test recommendations) were significantly more likely to be followed by EHR action; these alerts were also more likely to

prompt medication-specific action such as medication ordering, changing, and discontinuing or medication list viewing.

Not all alerts were followed by provider action, indicating that other factors may have played a role in providers' decision to act on information or recommendations provided in the alerts.

Data from access and audit logs provide a rare "behind-the-scenes" glimpse into how health care team members spend their time and where they direct their attention. These log data have been used successfully to understand questions of the health care team's communication [26], trainee skill [27], and clinical workflow [6,7]. As we demonstrate in this paper, EHR logs can also provide valuable implementation data that can be used to improve future CDS interventions. Log data can help a research team understand whether a CDS message was opened by the addressed recipient, how long it was viewed for, and what kind of actions followed the viewing.

There are numerous ways a research team can benefit from this knowledge. Our finding that few physician, patient, and contextual aspects impacted the opening of and action on alerts is reassuring. It does not appear that subgroups of PCPs or patients would require varying approaches for alert-based communication. Tracking rates of opening for noninterruptive messages can identify a need for improved safety measures (eg, escalation of unopened time-sensitive messages). Understanding who opens these messages lends insight into a care team's triage processes, which might alter the design of the intervention (eg, physician-directed CDS sent to a team where all messages are opened initially by nurses might prompt the team to modify message content or change the method of communication) [4]. Realizing that providers spend an extremely short period of time with alerts open might prompt those designing these alerts to adapt messages accordingly (eg, include fewer words and optimize visibility of key text). For the intervention discussed in this paper, we preceded implementation by having two of the group practice primary care physicians review all messages generated by the system for four months. Of the tested alert types, the reviewers selected those included in the final intervention with the goal that all alerts would be necessary and actionable. For the medication information-only alerts, which focused specifically on new or high-risk medications, the reviewers were certain that recipients would review the full medication list for these patients. Alerts that were categorized as information-only may have been perceived by PCPs as not requiring further action, since they did not include specific recommendations. The brief viewing time and low level of activity after viewing information-only alerts suggests that this type of alert should be reconsidered before including them in future interventions. Tracking patterns of EHR behavior after alert opening might suggest useful shortcuts to make the alert more user-friendly (eg, if users routinely navigate to the medication list after opening the alert, providing a link directly to this destination might enhance the user experience). In sum, to understand why an EHR-based CDS intervention is (or is not) effective, researchers should be asking whether that intervention was delivered as planned and what subsequent steps were taken by recipients after intervention delivery.

In this descriptive study, we focused on use of access and audit logs for understanding the impact of noninterruptive alerts. This approach is particularly necessary for studying noninterruptive alerts, which are usually opened outside of a patient encounter and thus present challenges for those seeking to pinpoint the timing of downstream actions. Access and audit log data can also be applied to a variety of questions pertaining to transmission of EHR-based information. Examples include tracking of information viewing (eg, understanding when a care team member opens and reads a clinic note, laboratory or test report, or other updated EHR information) as well as tracking the timing of documentation and ordering.

Access and audit logs could also be helpful in trying to reduce the InBasket burden for physicians. Log data can reveal which types of alerts and notifications physicians act upon; those that are commonly left unopened or are not generally followed by any action could be candidates for elimination or modification. These data can also inform enhancements to the structure of the InBasket. Unopened or unacted upon messages can be moved to and organized within separate folders to which the providers can return at a later time.

In addition to the two challenges identified here (difficulties in capturing a historical snapshot from the EHR and difficulty in interpretation of data elements in the absence of clear documentation), there are additional issues to consider when using access and audit log data. Site-to-site variability in the use of certain variables may introduce differences in interpretation of audit logs across sites. Data managers may not have privileges granted to retrieve access or audit log data. The large volume of data generated through access and audit logs may necessitate policies for limitations in what a site stores (and for how long). Understanding the local situation at an early stage of research planning may facilitate use (for instance, planning to download log data before they are deleted at a site that stores them for only a limited period). Planning ahead and setting up log data to provide real-time information as an intervention is implemented can make the process much simpler.

There are limitations to this study. We may not have captured all PCP actions after alert opening relevant to the postdischarge patient, such as actions taken outside of the EHR. The age of the data (from a trial conducted in 2011) may not accurately reflect current provider workflow in upgraded versions of the EHR with respect to opening of alerts and responsive actions. Direct causality (EHR action due to alert opening) is not proven simply by demonstrating close temporal proximity. Our study is descriptive and could be supported by future qualitative work examining the physicians' own perspectives on their EHR actions following alert opening.

An EHR-based alert system intervention alone does not improve clinical patient outcomes in high-risk populations. Nonetheless, the availability of information in the EHR can benefit similar studies trying to understand the link between provider behavior in the EHR and patient care and outcomes. Although analysis of EHR logs presents many challenges, data from these logs can provide researchers with insight into designing impactful EHR-based CDS and alert interventions.

Acknowledgments

This study was funded by a grant from the Agency for Healthcare Research and Quality (AHRQ) under award number 5R21HS023661. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH).

The views expressed in this article are those of the authors and do not represent the views of the US Department of Veterans Affairs or the United States Government.

Conflicts of Interest

None declared.

References

1. Boxwala AA, Kim J, Grillo JM, Ohno-Machado L. Using statistical and machine learning to help institutions detect suspicious access to electronic health records. *J Am Med Inform Assoc* 2011;18(4):498-505. [Medline: [21672912](#)]
2. Menon AK, Jiang X, Kim J, Vaidya J, Ohno-Machado L. Detecting Inappropriate Access to Electronic Health Records Using Collaborative Filtering. *Mach Learn* 2014 Apr 01;95(1):87-101 [FREE Full text] [doi: [10.1007/s10994-013-5376-1](#)] [Medline: [24683293](#)]
3. Freeman M. Miller Weisbrod. Dallas, TX: Miller Weisbrod, LLP Electronic Records and Audit Trails URL: https://www.millerweisbrod.com/docs/max/Electronic_Records_and_Audit_Trails.pdf [accessed 2019-01-30] [WebCite Cache ID [75pQHARw4](#)]
4. Cutrona SL, Fouayzi H, Burns L, Sadasivam RS, Mazor KM, Gurwitz JH, et al. Primary Care Providers' Opening of Time-Sensitive Alerts Sent to Commercial Electronic Health Record InBaskets. *J Gen Intern Med* 2017 Nov;32(11):1210-1219. [doi: [10.1007/s11606-017-4146-3](#)] [Medline: [28808942](#)]
5. Tai-Seale M, Olson CW, Li J, Chan AS, Morikawa C, Durbin M, et al. Electronic Health Record Logs Indicate That Physicians Split Time Evenly Between Seeing Patients And Desktop Medicine. *Health Aff (Millwood)* 2017 Dec 01;36(4):655-662 [FREE Full text] [doi: [10.1377/hlthaff.2016.0811](#)] [Medline: [28373331](#)]
6. Hirsch AG, Jones JB, Lerch VR, Tang X, Berger A, Clark DN, et al. The electronic health record audit file: the patient is waiting. *J Am Med Inform Assoc* 2017 Apr 01;24(e1):e28-e34. [doi: [10.1093/jamia/ocw088](#)] [Medline: [27375293](#)]
7. Hribar MR, Read-Brown S, Reznick L, Lombardi L, Parikh M, Yackel TR, et al. Secondary Use of EHR Timestamp data: Validation and Application for Workflow Optimization. *AMIA Annu Symp Proc* 2015;2015:1909-1917 [FREE Full text] [Medline: [26958290](#)]
8. Bowes WA. Measuring use of electronic health record functionality using system audit information. *Stud Health Technol Inform* 2010;160(Pt 1):86-90. [Medline: [20841655](#)]
9. Lo HG, Matheny ME, Seger DL, Bates DW, Gandhi TK. Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *J Am Med Inform Assoc* 2009;16(1):66-71 [FREE Full text] [doi: [10.1197/jamia.M2687](#)] [Medline: [18952945](#)]
10. Singh H, Thomas EJ, Mani S, Sittig D, Arora H, Espadas D, et al. Timely follow-up of abnormal diagnostic imaging test results in an outpatient setting: are electronic medical records achieving their potential? *Arch Intern Med* 2009 Sep 28;169(17):1578-1586 [FREE Full text] [doi: [10.1001/archinternmed.2009.263](#)] [Medline: [19786677](#)]
11. Singh H, Arora HS, Vij MS, Rao R, Khan MM, Petersen LA. Communication outcomes of critical imaging results in a computerized notification system. *J Am Med Inform Assoc* 2007;14(4):459-466 [FREE Full text] [doi: [10.1197/jamia.M2280](#)] [Medline: [17460135](#)]
12. Singh H, Thomas EJ, Sittig DF, Wilson L, Espadas D, Khan MM, et al. Notification of abnormal lab test results in an electronic medical record: do any safety concerns remain? *Am J Med* 2010 Mar;123(3):238-244 [FREE Full text] [doi: [10.1016/j.amjmed.2009.07.027](#)] [Medline: [20193832](#)]
13. Slovis BH, Nahass TA, Salmasian H, Kuperman G, Vawdrey DK. Asynchronous automated electronic laboratory result notifications: a systematic review. *J Am Med Inform Assoc* 2017 Nov 01;24(6):1173-1183. [doi: [10.1093/jamia/ocx047](#)] [Medline: [28520977](#)]
14. Dalal AK, Roy CL, Poon EG, Williams DH, Nolido N, Yoon C, et al. Impact of an automated email notification system for results of tests pending at discharge: a cluster-randomized controlled trial. *J Am Med Inform Assoc* 2014;21(3):473-480 [FREE Full text] [doi: [10.1136/amiajnl-2013-002030](#)] [Medline: [24154834](#)]
15. O'Connor SD, Dalal AK, Sahni VA, Lacson R, Khorasani R. Does integrating nonurgent, clinically significant radiology alerts within the electronic health record impact closed-loop communication and follow-up? *J Am Med Inform Assoc* 2016 Mar;23(2):333-338 [FREE Full text] [doi: [10.1093/jamia/ocv105](#)] [Medline: [26335982](#)]

16. Lacson R, O'Connor SD, Andriole KP, Prevedello LM, Khorasani R. Automated critical test result notification system: architecture, design, and assessment of provider satisfaction. *AJR Am J Roentgenol* 2014 Nov;203(5):W491-W496 [FREE Full text] [doi: [10.2214/AJR.14.13063](https://doi.org/10.2214/AJR.14.13063)] [Medline: [25341163](https://pubmed.ncbi.nlm.nih.gov/25341163/)]
17. Abujudeh HH, Kaewlai R, Choy G, Whelton DG, Rosenthal DI. Important imaging finding e-mail alert system: experience after 3 years of implementation. *Radiology* 2009 Sep;252(3):747-753. [doi: [10.1148/radiol.2531082217](https://doi.org/10.1148/radiol.2531082217)] [Medline: [19703870](https://pubmed.ncbi.nlm.nih.gov/19703870/)]
18. Benthin C, Pannu S, Khan A, Gong M, NHLBI Prevention Early Treatment of Acute Lung Injury (PETAL) Network. The Nature and Variability of Automated Practice Alerts Derived from Electronic Health Records in a U.S. Nationwide Critical Care Research Network. *Ann Am Thorac Soc* 2016 Dec;13(10):1784-1788 [FREE Full text] [doi: [10.1513/AnnalsATS.201603-172BC](https://doi.org/10.1513/AnnalsATS.201603-172BC)] [Medline: [27404288](https://pubmed.ncbi.nlm.nih.gov/27404288/)]
19. Gurwitz JH, Field TS, Ogarek J, Tjia J, Cutrona SL, Harrold LR, et al. An electronic health record-based intervention to increase follow-up office visits and decrease rehospitalization in older adults. *J Am Geriatr Soc* 2014 May;62(5):865-871 [FREE Full text] [doi: [10.1111/jgs.12798](https://doi.org/10.1111/jgs.12798)] [Medline: [24779524](https://pubmed.ncbi.nlm.nih.gov/24779524/)]
20. Murphy DR, Reis B, Sittig DF, Singh H. Notifications received by primary care practitioners in electronic health records: a taxonomy and time analysis. *Am J Med* 2012 Feb;125(2):209.e1-209.e7. [doi: [10.1016/j.amjmed.2011.07.029](https://doi.org/10.1016/j.amjmed.2011.07.029)] [Medline: [22269625](https://pubmed.ncbi.nlm.nih.gov/22269625/)]
21. Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. *JAMA Intern Med* 2013 Apr 22;173(8):702-704 [FREE Full text] [doi: [10.1001/2013.jamainternmed.61](https://doi.org/10.1001/2013.jamainternmed.61)] [Medline: [23460235](https://pubmed.ncbi.nlm.nih.gov/23460235/)]
22. Murphy DR, Meyer AND, Russo E, Sittig DF, Wei L, Singh H. The Burden of Inbox Notifications in Commercial Electronic Health Records. *JAMA Intern Med* 2016 Apr;176(4):559-560 [FREE Full text] [doi: [10.1001/jamainternmed.2016.0209](https://doi.org/10.1001/jamainternmed.2016.0209)] [Medline: [26974737](https://pubmed.ncbi.nlm.nih.gov/26974737/)]
23. Gregory ME, Russo E, Singh H. Electronic Health Record Alert-Related Workload as a Predictor of Burnout in Primary Care Providers. *Appl Clin Inform* 2017 Jul 05;8(3):686-697 [FREE Full text] [doi: [10.4338/ACI-2017-01-RA-0003](https://doi.org/10.4338/ACI-2017-01-RA-0003)] [Medline: [28678892](https://pubmed.ncbi.nlm.nih.gov/28678892/)]
24. Field TS, Garber L, Gagne SJ, Tjia J, Preusse P, Donovan JL, et al. Technological resources and personnel costs required to implement an automated alert system for ambulatory physicians when patients are discharged from hospitals to home. *Inform Prim Care* 2012;20(2):87-93 [FREE Full text] [Medline: [23710773](https://pubmed.ncbi.nlm.nih.gov/23710773/)]
25. Tjia J, Field TS, Garber LD, Donovan JL, Kanaan AO, Raebel MA, et al. Development and pilot testing of guidelines to monitor high-risk medications in the ambulatory setting. *Am J Manag Care* 2010 Jul;16(7):489-496 [FREE Full text] [Medline: [20645664](https://pubmed.ncbi.nlm.nih.gov/20645664/)]
26. Gray JE, Feldman H, Reti S, Markson L, Lu X, Davis RB, et al. Using Digital Crumbs from an Electronic Health Record to identify, study and improve health care teams. *AMIA Annu Symp Proc* 2011;2011:491-500 [FREE Full text] [Medline: [22195103](https://pubmed.ncbi.nlm.nih.gov/22195103/)]
27. Chen ES, Cimino JJ. Automated discovery of patient-specific clinician information needs using clinical information system log files. *AMIA Annu Symp Proc* 2003:145-149 [FREE Full text] [Medline: [14728151](https://pubmed.ncbi.nlm.nih.gov/14728151/)]

Abbreviations

- AOR:** adjusted odds ratio
- CDS:** clinical decision support
- EHR:** electronic health record
- N/A:** not applicable
- OR:** odds ratio
- PCP:** primary care physician

Edited by G Eysenbach; submitted 30.10.18; peer-reviewed by D Carvalho, D Murphy; comments to author 21.11.18; revised version received 14.01.19; accepted 20.01.19; published 07.02.19.

Please cite as:

Amroze A, Field TS, Fouayzi H, Sundaresan D, Burns L, Garber L, Sadasivam RS, Mazor KM, Gurwitz JH, Cutrona SL
Use of Electronic Health Record Access and Audit Logs to Identify Physician Actions Following Noninterruptive Alert Opening: Descriptive Study
JMIR Med Inform 2019;7(1):e12650
URL: <http://medinform.jmir.org/2019/1/e12650/>
doi: [10.2196/12650](https://doi.org/10.2196/12650)
PMID: [30730293](https://pubmed.ncbi.nlm.nih.gov/30730293/)

©Azraa Amroze, Terry S Field, Hassan Fouayzi, Devi Sundaresan, Laura Burns, Lawrence Garber, Rajani S Sadasivam, Kathleen M Mazor, Jerry H Gurwitz, Sarah L Cutrona. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 07.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Completeness of Electronic Dental Records in a Student Clinic: Retrospective Analysis

Seth Aaron Levitin^{1*}, BSc; John T Grbic¹, DMD; Joseph Finkelstein^{2*}, MD, PhD

¹Division of Foundational Sciences, Columbia University College of Dental Medicine, New York, NY, United States

²Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, United States

*these authors contributed equally

Corresponding Author:

Joseph Finkelstein, MD, PhD

Department of Population Health Science and Policy

Icahn School of Medicine at Mount Sinai

1425 Madison Ave, Rm L2-36

New York, NY, 10029

United States

Phone: 1 212 659 9596

Email: Joseph.Finkelstein@mssm.edu

Abstract

Background: A well-designed, adequately documented, and properly maintained patient record is an important tool for quality assurance and care continuity. Good clinical documentation skills are supposed to be a fundamental part of dental student training.

Objective: The goal of this study was to assess the completeness of electronic patient records in a student clinic.

Methods: Completeness of patient records was assessed using comparative review of validated cases of alveolar osteitis treated between August 2011 and May 2017 in a student clinic at Columbia University College of Dental Medicine, New York, USA. Based on a literature review, population-based prevalence of nine most frequently mentioned symptoms, signs, and treatment procedures of alveolar osteitis was identified. Completeness of alveolar osteitis records was assessed by comparison of population-based prevalence and frequency of corresponding items in the student documentation. To obtain all alveolar osteitis cases, we ran a query on the electronic dental record, which included all cases with diagnostic code Z1820 or any variation of the phrases “dry socket” and “alveolar osteitis” in the notes. The resulting records were manually reviewed to definitively confirm alveolar osteitis and to extract all index items.

Results: Overall, 296 definitive cases of alveolar osteitis were identified. Only 22% (64/296) of cases contained a diagnostic code. Comparison of the frequency of the nine index categories in the validated alveolar osteitis cases between the student clinic and the population showed the following results: severe pain: 94% (279/296) vs 100% (430/430); bare bone/missing blood clot: 27% (80/296) vs 74% (35/47) to 100% (329/329); malodor: 7% (22/296) vs 33%-50% (18/54); radiating pain to the ear: 8% (24/296) vs 56% (30/54); lymphadenopathy: 1% (3/296) vs 9% (5/54); inflammation: 14% (42/296) vs 50% (27/54); debris: 12% (36/296) vs 87% (47/54); alveolar osteitis site noted: 96% (283/296) vs 100% (430/430; accepted documentation requirement); and anesthesia during debridement: 77% (20/24) vs 100% (430/430; standard of anesthetization prior to debridement).

Conclusions: There was a significant discrepancy between the index category frequency in alveolar osteitis cases documented by dental students and in the population (reported in peer-reviewed literature). More attention to clinical documentation skills is warranted in dental student training.

(*JMIR Med Inform* 2019;7(1):e13008) doi:[10.2196/13008](https://doi.org/10.2196/13008)

KEYWORDS

electronic medical records; patient record completeness; dentistry

Introduction

A clinical record is a fundamental part of patient care delivery [1]. Its completeness is important for many reasons. A record's

main purpose is to serve as a means of communication among providers themselves and between providers and their colleagues [2]. Clinical decision support is dependent on accurate and complete dental records [3], which also aid in the evaluation of

a patient's care [4]. In the event of a lawsuit, the record serves as evidence [5] and its contents are necessary to determine whether the diagnosis and treatment met appropriate standards [6]. With the expansion of dental informatics applications, it is even more essential to have a complete record in order to ensure proper analysis and results in outcomes research [5]. However, there are known issues with the completeness of electronic medical records (EMRs).

In the past decade, electronic patient records became a ubiquitous part of dental care delivery [7]. Oral health data accumulated in the process of clinical care represent a rich and readily available recourse for scientific investigation and data analytics [8]. Recent analysis of electronic dental records (EDRs) helped identify predictors of implant survival [9] as well as the prevalence and risk factors of peri-implantitis [10]. Application of machine learning techniques [11] and temporal analytics [12] resulted in new opportunities for knowledge discovery and predictive analytics [13]. With the increasing use of EMRs as an important resource for scientific discovery, potential barriers for secondary analysis of EMR have been recognized [14]. Completeness of EMR is one of the most frequently discussed issues that may limit the use of EMR data for clinical and population health research [15]. This issue was reported to be particularly relevant in the evaluation of medical student documentation [16]; however, no systematic assessment of the completeness of EMR in dental student clinics has been performed.

Since the widespread introduction of EMRs, numerous studies have indicated gaps in documentation. A study at the University of Michigan examined whether there were differences in the reported eye symptoms between EMRs and eye symptom questionnaires that patients fill out. Exact agreement was found in only 23.5% (38/162) of cases. In cases where patients reported three or more symptoms, data from the eye symptom questionnaire always varied from the EMR data [17]. A similar study at the Mayo Clinic compared symptoms of chest pain, dyspnea, and cough between information forms patients received prior to the appointment and EMRs, for the purpose of identifying stable angina pectoris. They found that the two documents had varying levels of positive agreement (ratio): 74 for chest pain, 70 for dyspnea, and 63 for cough [18]. Researchers at Duke University analyzed records for completeness for quality purposes, but concluded that improper documentation for colorectal cancer impeded their ability to accurately calculate patient performance. Of the 499 patients eligible for the analysis, only 66 had sufficient documentation. In addition, only 86% (427/499) of EMRs confirmed a diagnosis, 29% (143/499) were missing the age, and only 38% (188/499) stated the TMN stage [19].

Other studies indicated that the quality of documents in EDRs may be suboptimal [5,20]. However, there is a lack of systematic studies on the completeness of EDRs. The goal of this project was to review the documentation quality of EDRs in a dental school clinic.

Methods

Data Source

EDRs of patients examined at Columbia University's College of Dental Medicine were analyzed.

Data Collection

The study was a retrospective analysis of EDRs from patients with dry sockets diagnosed at Columbia University's College of Dental Medicine between August 2011 and May 2017. We ran a query on the college's database to find EDRs containing diagnostic code Z1820, the phrases "dry socket" and "alveolar osteitis," or a variation of those two phrases. All queries were performed using structured query language in the Oracle database containing data from an EDR system called axiUm (Exan Group, Las Vegas, NV).

For the purpose of this study, a dry socket was defined by a diagnostic Z code at the initial encounter. Alternatively, it was defined by the presence of a key word in a note combined with clinical evidence and explicit documentation of the patient with alveolar osteitis or dry socket. Z codes are represented by a list of diagnostic codes and terms developed for use with EDR, as previously described [21].

After reviewing the literature on dry socket [22-28], we compiled a list of 17 criteria related to dry socket. *Five* were related to treatment: curettage, irrigation, anesthesia, intra-alveolar medication, and medication. *Three* were symptoms: pain, radiating pain, and tenderness on palpation. *Nine* were signs: lack of blood clot, malodor, low-grade fever, bare bone, lymphadenopathy, pus, erythema, inflamed gingiva/socket, and debris. The *last* category was the socket site of dry socket. Severe pain was a necessary symptom for dry socket diagnosis [25,28].

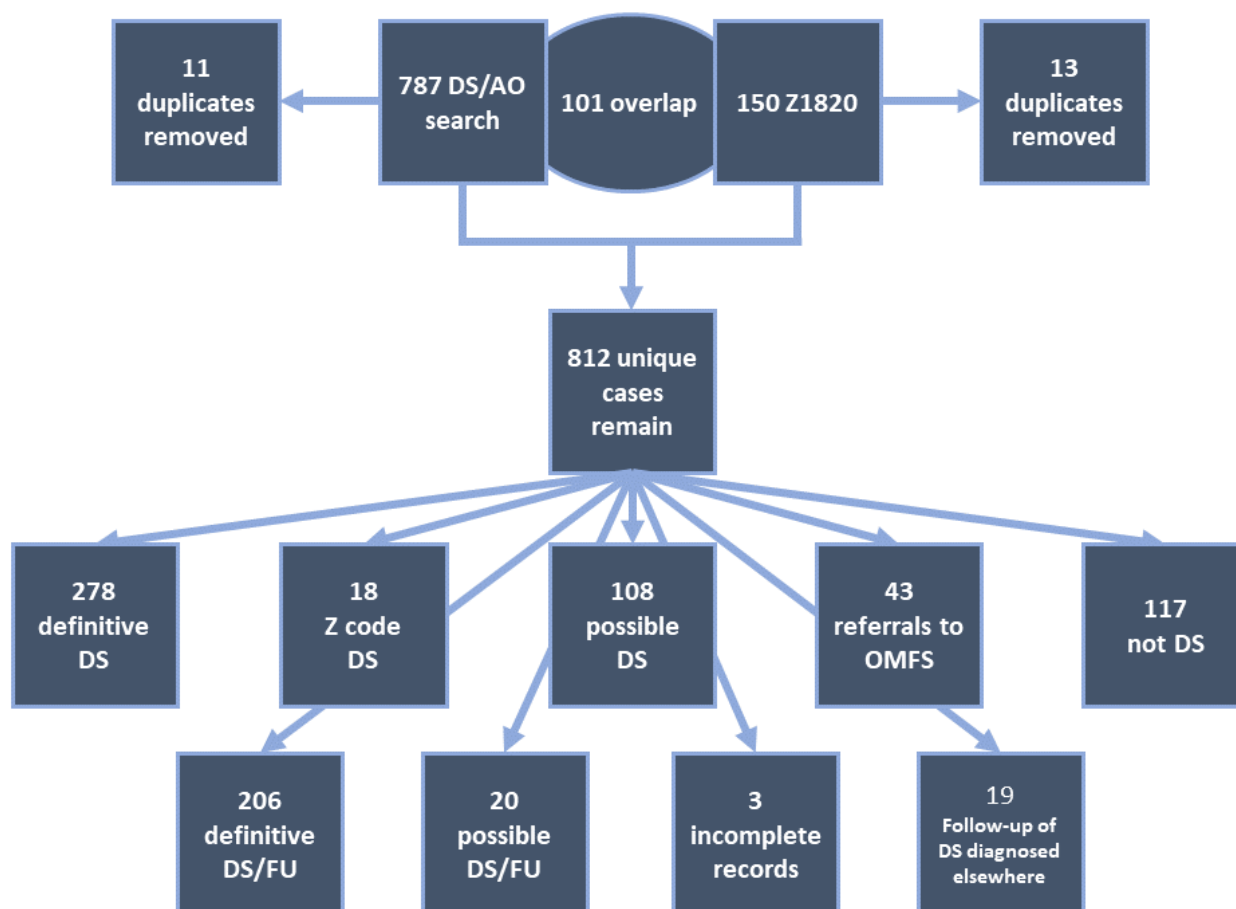
Data Analysis

A dental student reviewed each of the query results to confirm the presence of a dry socket. The cases were also analyzed for the presence of any of the 17 criteria mentioned above. Pertinent positive and negative results for dry socket criteria were both recorded for further analysis. The student then reviewed the available literature on dry socket statistics to determine the prevalence of symptoms in the literature. Both the positive and negative criteria were compared to the baseline figures from this literature.

Results

A total of 150 records with diagnostic Z codes were identified, and another 787 records were identified by searching for a mention of dry socket or alveolar osteitis in the notes (Figure 1). Both queries resulted in a number of duplicate cases, which were removed. The dry socket/alveolar osteitis dataset had 11 duplicate cases, and the Z code data set contained 13 duplicate cases. In addition, 101 of the remaining cases overlapped (ie, had both diagnostic Z codes *and* dry socket or alveolar osteitis in the notes). These overlapping cases were not counted twice in the study. After removing duplicates and overlapping cases, the final number of combined cases was 812.

Figure 1. Breakdown of the reviewed cases. DS: dry socket; AO: alveolar osteitis; FU: follow-up; OMFS: Department of Oral and Maxillofacial Surgery.



Cases were then categorized based on the dental record content (Table 1). An explicit diagnosis of dry socket was found in 278 EMRs. Another 18 EMRs had a diagnostic Z code and the patient received treatment for dry sockets, but the dentist never explicitly stated that the patient presented with a dry socket; these were also included as definitive dry sockets for the purpose of this study. Totally, 296 cases (of 812, 36.5%) were categorized as definitive dry sockets, and 108 cases (of 812, 13.3%) of possible dry sockets had notes of treatment of dry sockets, but the dentist did not include a diagnostic code or any definitive diagnoses of dry socket in the note. This misdocumentation is important, but could not be included in the data, as there is no way to confirm the patient had a dry socket from the records. The remaining cases were classified as follows:

- Follow-up treatment of both definitive dry sockets and possible dry sockets (226/812, 27.8%).
- Cases that were definitively classified as nondry sockets: These patients' records were included in the query because they contained the key word to rule out the diagnosis, history of previous dry socket, mention of dry socket paste, or similar reference (117/812, 14.4%).
- Referrals to the oral surgery clinic: Such referrals for further evaluation were also a common finding. The referrals were not followed up to determine if these patients were later seen in Columbia (43/812, 5.3%). An additional segment

of charts belonged to patients whose only encounter populated in our search was for a follow-up in our clinic for a dry socket (19/812, 2.3%). The final and smallest group of charts contained a Z code, but no note was found in the system for the encounter. This may have been a kink in the system or oversight by the provider (3/812, 0.4%).

Definitive dry socket cases were analyzed for correctness of EMR documentation. The demographics of patients of the 296 records are outlined in Table 2. Women (196/296, 66.3%) were affected by dry socket more frequently than men. The age group with the highest frequency of dry socket included adults between the ages of 20 and 39 years (146/296, 49.3%). Table 3 demonstrates the extraction characteristics of the definitive dry sockets studied. Molars were the most frequent sites for dry socket, with the third molar represented in 37.5% of the cases (111/296). Tooth extraction was performed under the supervision of faculty from the Department of Oral and Maxillofacial Surgery in 82.8% (245/296) cases. About half of the tooth extraction procedures were carried out by predoctoral students (145/296, 46.3%) and the remaining extractions were performed by postdoctoral students, the majority of whom were represented by residents of the Department of Oral and Maxillofacial Surgery. A similar pattern was observed for location and provider type for the tooth extraction follow-up visits during which dry socket was diagnosed and treated.

Table 1. Categories of classifications.

Category	Description	n
Definitive dry socket	Diagnostic code with signs, symptoms, and treatment, or explicit clinical diagnosis	296
Dry socket follow-up	Follow-up to case classified as definitive dry socket, possible dry socket, or other treatment	230
Possible dry socket	Presented with symptoms and treated for dry socket, but no explicit diagnosis or Z code	108
Referral	Referred to the Department of Oral and Maxillofacial Surgery	43
Not dry socket	Diagnosis excluded by provider, history of previous dry socket in notes, or similar finding	117
Dry socket treatment follow-up or missing other documentation	Only the follow-up on previous treatment was contained in the query	15
Incomplete notes	Notes were not populated but contained the Z code	3

Table 2. Patient demographics.

Characteristic	n (%)
Gender	
Male	100 (33.7)
Female	196 (66.3)
Age (years)	
<20	11 (3.7)
20-39	146 (49.3)
40-59	82 (27.7)
60-79	53 (17.9)
>80	4 (1.4)
Ethnicity	
African American	23 (7.8)
Asian	4 (1.4)
Caucasian	9 (3.0)
Hispanic	120 (40.5)
Other	139 (47.0)
Not disclosed	1 (0.3)
Language	
English	195 (65.88)
Spanish	68 (22.97)
Other	30 (10.14)
Arabic	2 (0.68)
Russian	1 (0.34)

Two general types of misdocumentation were encountered and calculated while analyzing the 296 cases of definitive dry socket (Table 4). The first was any missing gross documentation necessary for diagnosis of dry socket to support the clinical diagnosis of this condition and to provide documentation necessary for patient follow-up and care continuity. These included documented pain, visible bone, or lack of blood clot. We also included the missing socket position when diagnosing

a dry socket. Of the 296 cases, 220 (74.3%) were missing at least one of the abovementioned factors. The second type of misdocumentation was the number of cases without a diagnostic Z code. Of the 296 cases, 232 (78.4%) did not contain the diagnostic code. All cases analyzed were documented after the Columbia University College of Dental Medicine implemented diagnostic codes.

Table 3. Extraction characteristics.

Characteristics	n (%)
Extracted tooth	
Third molar	111 (37.5)
First and second molars	125 (42.2)
Premolars	40 (13.5)
Department of extraction treatment	
Department of Oral and Maxillofacial Surgery	245 (82.8)
Periodontics	13 (4.4)
Other	38 (12.8)
Provider group for extraction	
Student	145 (46.3)
Department of Oral and Maxillofacial Surgery	65 (24.1)
General practice residency	23 (8.5)
Advanced education in general dentistry	18 (6.7)
Periodontics	14 (5.2)
Other	31 (10.4)
Location of dry socket treatment	
Department of Oral and Maxillofacial Surgery	242 (81.8)
Advanced education in general dentistry	17 (5.7)
Periodontics	14 (4.7)
Other	23 (7.8)
Provider groups for dry socket diagnosis	
Student	143 (48.3)
Department of Oral and Maxillofacial Surgery	49 (16.55)
General practice residency	42 (14.19)
Advanced education in general dentistry	26 (8.78)
Periodontics	13 (4.39)
Other	23 (7.8)

Possible misdocumentation was also found in the remaining 516 cases that were not labeled as definitive dry socket. There were 108 cases of patients who presented with symptoms of dry sockets and received treatment for dry sockets. However, their charts did not include a Z code or any definitive diagnoses by the provider. Although many of these patients fulfilled the requirements to be classified with a dry socket, they could not be included in the analysis because they lacked a basic documented confirmation. As previously stated, 206 cases were classified as follow-up appointments to initial treatment of definitive dry socket. From these cases, another category of misdocumentation emerged: 16 of the cases contained a diagnostic Z code, although the actual diagnosis and initial treatment of the dry socket did not contain this code. The latter

type of misdocumentation does not have any direct consequences, but reveals a reluctance to use dry socket diagnosis or lack of education regarding diagnostic codes.

The final group of misdocumentation was found in the 19 cases categorized as follow-up to the initial appointment for dry socket, in which the patient's first documented encounter with the word dry socket or alveolar osteitis was in the follow-up. The clinical note in 14 of these 19 cases indicated that the patient was previously seen in the clinic or emergency room to treat the dry socket. The implication is that these patients were seen and treated for a dry socket previously in the clinic, but the earlier documentation lacked any of the three search terms (dry socket, alveolar osteitis, and Z code) used to identify cases for this analysis.

Table 4. Types of misdocumentation encountered (N=296).

Type of misdocumentation	n (%)
Missing documentation necessary for diagnosis (pain, bone visible, or socket position)	220 (74.3)
No Z codes	232 (78.4)
Treated for dry socket and had symptoms, but no definitive diagnosis	108 (36.5) ^a
Dry socket follow-up is Z code but no initial encounter	16 (5.4) ^a
Dry socket follow-up was searchable but not the original case	14 (4.7) ^{a,b}

^aThese cases were categorized for purposes of misdocumentation but were not included in the analysis of misdocumentation of definitive dry sockets.

^bIncludes 2 from the emergency room.

Definitive dry socket signs, symptoms, and treatments completeness compared to the best values available in the literature are presented in Figure 2. Presence of severe pain was considered a necessary symptom for dry socket diagnosis. The clinical notes only contained pain as a symptom in 94% (279/296) of the cases. Bare bone [23], open socket [25], or missing blood clot [26] was mentioned in 74% (35/47) to 100% (329/329) of the dry socket cases. In our clinic, any of these related terms were mentioned in 27% (80/296) of cases. In the 296 definitive dry socket cases, the word “bone” was only found in 50 cases and “clot” was found only 25 times; some of these cases overlapped. These terms are germane and necessary for a diagnosis of dry socket, and the lack of documentation is troublesome. Malodor [22] was documented in 33% (18/54) to >50% of the dry socket cases in the literature [27]; however, its documentation was only present in 7% (22/296) of our cases. Radiating pain toward the ear was present in 56% (30/54) of the cases in the literature [22], but was mentioned only in 7% (24/296) of our cases (documentation of radiating pain alone

was included in this number, as we did not require ear or side of face to be documented). Lymphadenopathy [22] was present in 9% (5/54) of the cases in the literature [23], but in only 1% (3/296) of our cases. Inflammation was present in 50% (27/54) of the cases in the literature [22]; however, in the dental record, only 14% (42/296) of cases mentioned inflammation and 15% (44/296) (with some overlap with inflammation cases) mentioned erythema.

A comparison between the notes of predoctoral dental students and postdoctoral faculty and residents is outlined in Table 5. The nine selected signs, symptoms, and other documentation related to dry socket diagnosis were analyzed. Predoctoral students were more likely to properly document the location of the dry socket, while postdoctoral students were more likely to document malodor, presence of debris, and anesthesia administration during socket debridement. Misdocumentation of the remaining categories was equal (both groups were within 1% of one another for the other five categories) as compared to the expected values (Figure 2).

Figure 2. Signs and symptoms in misdocumentation. DS: dry socket; CDM: clinical dental record.

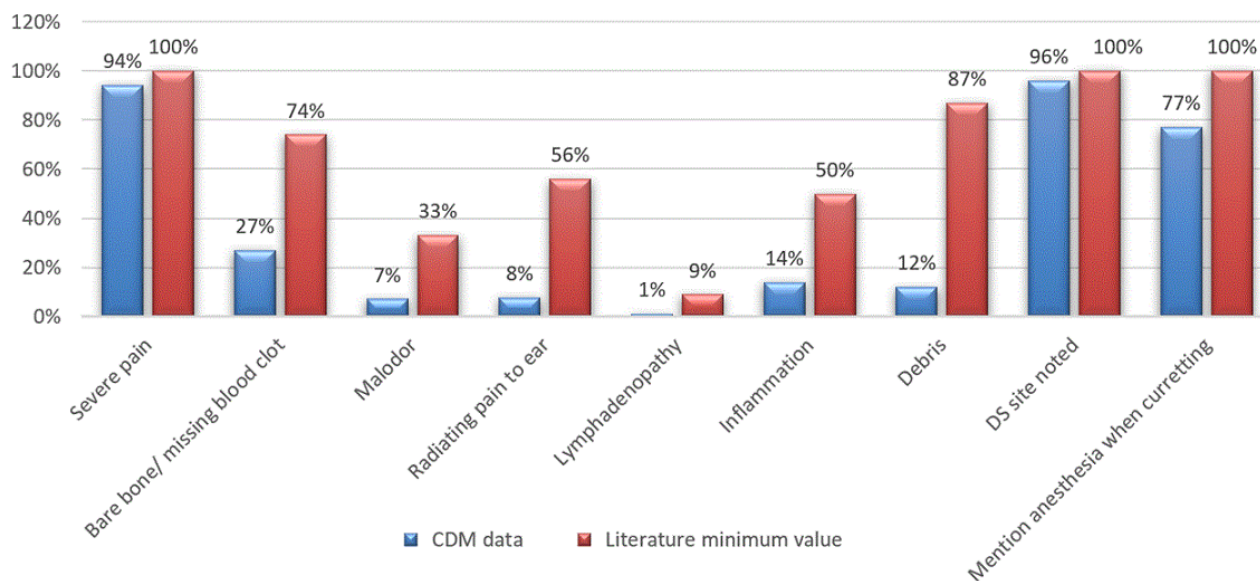


Table 5. Comparison between the notes of predoctoral dental students and postdoctoral faculty and residents.

Characteristic	Predoorate (n=143), n (%)	Postdoctorate (n=151), n (%)
Severe pain	135 (95)	144 (94)
Bare bone/missing blood clot	31 (22)	49 (32)
Malodor	6 (4)	16 (10)
Radiating pain to ear	12 (8)	12 (8)
Lymphadenopathy	1 (1)	2 (1)
Inflammation	21 (15)	21 (14)
Debris	12 (8)	24 (16)
Dry socket site noted	140 (98)	143 (93)
Anesthesia mentioned during debridement	2 ^a (50)	18 ^b (82)

^aTotal for this group is 4 students.

^bTotal for this group is 22 faculty and residents.

Discussion

Our ultimate finding was that misdocumentation occurs in EDRs to varying degrees within a dental school clinic. The dental records reviewed lack many signs and symptoms that are necessary and expected to be recorded for a diagnosis of dry socket. This misdocumentation was prevalent in the notes of both predoctoral and postdoctoral students. The major limitation of this study, as with other electronic record retrospective studies, was our ability to confirm the diagnosis of dry socket. We also could not confirm whether patients actually had the symptoms omitted. We were forced to rely on the information provided by the documenter.

Evaluation of completeness of symptom documentation in this study followed recent guidance on EMR data quality–assessment methodology [29] that promulgates the use of validated population-based prevalence as a gold standard [30–31]. Following these guidelines, the comparison between the validated symptom prevalence and symptom frequency found in our EDR review was used to identify completeness of symptom documentation in EDR in this study. This approach has been successfully used in a number of previous studies to identify completeness of smoking status records [32], obesity reporting [33], hypertension records [34], and depression prevalence [35]. To minimize bias toward inflating misdocumentation rates, whenever several verified estimates of a population-based frequency for a particular symptom were available, a comparison was made between the lowest population-based frequency and the frequency found in the EDR, as previously described [29–32].

Our findings of the demographics and tooth extraction characteristics among patients with dry socket are congruent with previous reports. As previously determined, women are more likely to have dry socket than men [25,36]. The tooth position is most likely to be a third molar compared to any other individual area [37]. People in their 20s and 30s are at a higher risk for alveolar osteitis [38]. Two major types of misdocumentation that were found and analyzed in this study have been previously mentioned in the literature, including lack

of supporting documentation for clinical diagnosis [25,28] and absence of appropriate diagnostic codes [39].

Analysis of characteristics of providers who performed tooth extraction and dry socket diagnosis confirmed the external validity [40] of our study sample, as the resulting characteristics accurately reflected routine dental care delivery patterns for these types of procedures occurring in the student clinic. The majority of the cases were carried out under supervision of a preceptor from the Department of Oral and Maxillofacial Surgery. About half of the extractions and dry socket diagnoses were carried out by predoctoral students, and the rest were performed by postdoctoral residents of whom approximately half were represented by residents of the Department of Oral and Maxillofacial Surgery. The fact that the predoctoral and postdoctoral dental surgeons were equally presented in our analysis supported the unbiased comparison between documentation quality of predoctoral and postdoctoral students.

The results of documentation completeness in EDRs in this study are corroborated by the following reports. In Minnesota, a discrepancy was found between the American Dental Association's recommendation for dental record accuracy and the actual accuracy in dental practices [41]. In Finland, researchers observed a discrepancy between the quality of treatment a dentist believed he/she provided and the treatment the patient actually received, as contained in the EDR [42]. EMRs have also been shown to have issues with accuracy [43]. In an adult cardiology clinic, researchers discovered “very poor” completeness values for signs such as chest pain and shortness of breath [44]. In a systemic review of EMR completeness in primary care, Thiru [45] found that records of diagnoses with clear clinical criteria had a higher rate of completeness than those without clear criteria. This is relevant to dry socket, a diagnosis with unclear criteria. A trauma center study also found incompleteness of certain categories in the EMR [46]. Similarly, a study conducted with inpatient records at Menelik II Referral Hospital, a government hospital in Addis Ababa, Ethiopia, found “low” EMR completeness compared to the expected standard of 100% [47]. Legal Medical Record Standards stated that “Each Medical Record shall contain sufficient, accurate information to identify the patient, support the diagnosis, justify the

treatment, document the course and results, and promote continuity of care among health care providers” [48]. Following this simple instruction can vastly improve the delivery of care.

To resolve the issues with incomplete or inaccurate records, dental education should emphasize more on proper documentation and ensure its incorporation into the clinic routine. Documenters need to remember that their records are not simply for their own convenience, but may serve legal, research-related, or forensic purposes [1,49]. To this end, each record must be complete, without implicit assumptions, and follow a method that makes it easily accessible to any reader. Thierer [50] found an improvement in EDR accuracy by incorporating an in-service intervention for faculty members and a Moodle site course on documentation for students.

Our findings have important implications for future research that uses EMR data. Better understanding of the potential limitations of electronic health record data use promotes fidelity and reproducibility of secondary data analysis [51]. A variety of approaches are being implemented to address the potential limitations of EMR data [52] such as deep learning techniques [53] for imputing missing data, symbolic operations for time interval analytics [54], and calibration to reduce measurement error in prevalence estimates based on EMR data [55]. A growing number of studies employ common data models combined with cross-linked semantic ontologies to harmonize

EMR data [56] and confirm with the Findability, Accessibility, Interoperability, and Reusability principles [57].

We believe there are three sequential steps necessary for improving EDRs. The first step is additional training predoctoral and postdoctoral students on the importance of note comprehensiveness. As per a systemic review by the Accreditation Council for Continuing Medical Education, a live intervention with interactive techniques is the most effective way to change a physician’s behavior to influence patient outcomes [58]. Interventions should therefore be constructed accordingly. The next step is adding disease- and condition-specific worksheets to the EDR. The worksheets should contain categories pertinent to the specific diagnosis, with drop-down boxes for the practitioner to complete. This prevents the inadvertent omission of crucial categories. The last step is adding a clinical decision support tool to EDRs. The tool embodies evidence-based dentistry, an approach being adopted by an increasing number of dental schools and practitioners. However, this tool operates properly only if practitioners enter complete and accurate data into EDRs in a way that computers can easily analyze. Thus, strict compliance with the first two steps is critical. This will result in refined EDRs, which can potentially lead to superior and safer delivery of care at a lower cost [59]. Although some information in EDRs may seem largely irrelevant, the EDR is a critical depository of data, with limitless research possibilities. If properly executed, it may improve diagnoses, treatment, and dentistry as a whole.

Conflicts of Interest

None declared.

References

- Devadiga A. What's the deal with dental records for practicing dentists? Importance in general and forensic dentistry. *J Forensic Dent Sci* 2014 Jan;6(1):9-15 [FREE Full text] [doi: [10.4103/0975-1475.127764](https://doi.org/10.4103/0975-1475.127764)] [Medline: [24695875](https://pubmed.ncbi.nlm.nih.gov/24695875/)]
- Ho Y, Gadd CS, Kohorst KL, Rosenbloom ST. A qualitative analysis evaluating the purposes and practices of clinical documentation. *Appl Clin Inform* 2014 Dec;5(1):153-168 [FREE Full text] [doi: [10.4338/ACI-2013-10-RA-0081](https://doi.org/10.4338/ACI-2013-10-RA-0081)] [Medline: [24734130](https://pubmed.ncbi.nlm.nih.gov/24734130/)]
- Wright A, Phansalkar S, Bloomrosen M, Jenders RA, Bobb AM, Halamka JD, et al. Best practices in clinical decision support: the case of preventive care reminders. *Appl Clin Inform* 2010 Dec;1(3):331-345 [FREE Full text] [doi: [10.4338/ACI-2010-05-RA-0031](https://doi.org/10.4338/ACI-2010-05-RA-0031)] [Medline: [21991299](https://pubmed.ncbi.nlm.nih.gov/21991299/)]
- American Dental Association. 2010. Dental Records URL: http://www.ada.org/~media/ADA/Public%20Programs/Files/MPRG_Dental_Records.pdf?la=en [accessed 2019-03-18] [WebCite Cache ID 74XuEt0nV]
- Mangalmurti S, Murtagh L, Mello M. Medical malpractice liability in the age of electronic health records. *N Engl J Med* 2010 Nov 18;363(21):2060-2067. [doi: [10.1056/NEJMHle1005210](https://doi.org/10.1056/NEJMHle1005210)] [Medline: [21083393](https://pubmed.ncbi.nlm.nih.gov/21083393/)]
- Acharya A, Shimpi N, Mahnke A, Mathias R, Ye Z. Medical care providers' perspectives on dental information needs in electronic health records. *J Am Dent Assoc* 2017 Dec;148(5):328-337. [doi: [10.1016/j.adaj.2017.01.026](https://doi.org/10.1016/j.adaj.2017.01.026)] [Medline: [28284418](https://pubmed.ncbi.nlm.nih.gov/28284418/)]
- Acharya A, Schroeder D, Schwei K, Chyou P. Update on electronic dental record and clinical computing adoption among dental practices in the United States. *Clin Med Res* 2017 Dec;15(3-4):59-74 [FREE Full text] [doi: [10.3121/cmr.2017.1380](https://doi.org/10.3121/cmr.2017.1380)] [Medline: [29229631](https://pubmed.ncbi.nlm.nih.gov/29229631/)]
- Gordon S, Camargo G, Mejia G, Sutherland J. Use of the dental electronic health record for research: Assessing demographic and oral health characteristics data for clinic patients. *J Dent Educ* 2018 Dec;82(12):1249-1257. [doi: [10.21815/JDE.018.130](https://doi.org/10.21815/JDE.018.130)] [Medline: [30504461](https://pubmed.ncbi.nlm.nih.gov/30504461/)]
- Hickin M, Shariff J, Jennette P, Finkelstein J, Papapanou P. Incidence and Determinants of Dental Implant Failure: A Review of Electronic Health Records in a U.S. Dental School. *J Dent Educ* 2017 Oct;81(10):1233-1242 [FREE Full text] [doi: [10.21815/JDE.017.080](https://doi.org/10.21815/JDE.017.080)] [Medline: [28966189](https://pubmed.ncbi.nlm.nih.gov/28966189/)]

10. Kordbacheh Changi K, Finkelstein J, Papapanou P. Peri-implantitis prevalence, incidence rate, and risk factors: A study of electronic health records at a U.S. dental school. *Clin Oral Implants Res* 2019 Feb 15. [doi: [10.1111/clr.13416](https://doi.org/10.1111/clr.13416)] [Medline: [30768875](https://pubmed.ncbi.nlm.nih.gov/30768875/)]
11. Yoon S, Choi T, Odlum M, Mitchell D, Kronish I, Davidson K, et al. Machine Learning to Identify Behavioral Determinants of Oral Health in Inner City Older Hispanic Adults. *Stud Health Technol Inform* 2018;251:253-256 [FREE Full text] [Medline: [29968651](https://pubmed.ncbi.nlm.nih.gov/29968651/)]
12. Shknevsky A, Shahar Y, Moskovitch R. Consistent discovery of frequent interval-based temporal patterns in chronic patients' data. *J Biomed Inform* 2017 Nov;75:83-95 [FREE Full text] [doi: [10.1016/j.jbi.2017.10.002](https://doi.org/10.1016/j.jbi.2017.10.002)] [Medline: [28987378](https://pubmed.ncbi.nlm.nih.gov/28987378/)]
13. Wu J, Roy J, Stewart W. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. *Med Care* 2010 Jun;48(6 Suppl):S106-S113. [doi: [10.1097/MLR.0b013e3181de9e17](https://doi.org/10.1097/MLR.0b013e3181de9e17)] [Medline: [20473190](https://pubmed.ncbi.nlm.nih.gov/20473190/)]
14. Verheij R, Curcin V, Delaney B, McGilchrist M. Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. *J Med Internet Res* 2018 May 29;20(5):e185 [FREE Full text] [doi: [10.2196/jmir.9134](https://doi.org/10.2196/jmir.9134)] [Medline: [29844010](https://pubmed.ncbi.nlm.nih.gov/29844010/)]
15. Weiskopf N, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013 Oct;46(5):830-836 [FREE Full text] [doi: [10.1016/j.jbi.2013.06.010](https://doi.org/10.1016/j.jbi.2013.06.010)] [Medline: [23820016](https://pubmed.ncbi.nlm.nih.gov/23820016/)]
16. Seo J, Kong H, Im S, Roh H, Kim D, Bae H, et al. A pilot study on the evaluation of medical student documentation: assessment of SOAP notes. *Korean J Med Educ* 2016 Jun;28(2):237-241 [FREE Full text] [doi: [10.3946/kjme.2016.26](https://doi.org/10.3946/kjme.2016.26)] [Medline: [26996436](https://pubmed.ncbi.nlm.nih.gov/26996436/)]
17. Valikodath NG, Newman-Casey PA, Lee PP, Musch DC, Niziol LM, Woodward MA. Agreement of Ocular Symptom Reporting Between Patient-Reported Outcomes and Medical Records. *JAMA Ophthalmol* 2017 Mar 01;135(3):225-231 [FREE Full text] [doi: [10.1001/jamaophthalmol.2016.5551](https://doi.org/10.1001/jamaophthalmol.2016.5551)] [Medline: [28125754](https://pubmed.ncbi.nlm.nih.gov/28125754/)]
18. Pakhomov S, Jacobsen S, Chute C, Roger V. Agreement between patient-reported symptoms and their documentation in the medical record. *Am J Manag Care* 2008 Aug;14(8):530-539 [FREE Full text] [Medline: [18690769](https://pubmed.ncbi.nlm.nih.gov/18690769/)]
19. Abernethy AP, Herndon JE, Wheeler JL, Rowe K, Marcello J, Patwardhan M. Poor documentation prevents adequate assessment of quality metrics in colorectal cancer. *J Oncol Pract* 2009 Jul;5(4):167-174 [FREE Full text] [doi: [10.1200/JOP.0942003](https://doi.org/10.1200/JOP.0942003)] [Medline: [20856630](https://pubmed.ncbi.nlm.nih.gov/20856630/)]
20. Kopsala JT. WiredSpace. 2014. Quality Assurance and Dental Patient Records URL: <http://wiredspace.wits.ac.za/handle/10539/14625> [accessed 2019-03-18] [WebCite Cache ID 74XzZMKiY]
21. White JM, Kalenderian E, Stark PC, Ramoni RL, Vaderhobli R, Walji MF. Evaluating a dental diagnostic terminology in an electronic health record. *J Dent Educ* 2011 May;75(5):605-615 [FREE Full text] [Medline: [21546594](https://pubmed.ncbi.nlm.nih.gov/21546594/)]
22. Singh V, Pal US, Singh R, Soni N. Honey a sweet approach to alveolar osteitis: A study. *Natl J Maxillofac Surg* 2014 Jan;5(1):31-34 [FREE Full text] [doi: [10.4103/0975-5950.140166](https://doi.org/10.4103/0975-5950.140166)] [Medline: [25298714](https://pubmed.ncbi.nlm.nih.gov/25298714/)]
23. Soni N, Singh V, Mohammad S, Singh RK, Pal US, Singh R, et al. Effects of honey in the management of alveolar osteitis: A study. *Natl J Maxillofac Surg* 2016;7(2):136-147 [FREE Full text] [doi: [10.4103/0975-5950.201354](https://doi.org/10.4103/0975-5950.201354)] [Medline: [28356684](https://pubmed.ncbi.nlm.nih.gov/28356684/)]
24. Baqain Z, Karaky AA, Sawair F, Khraisat A, Khaisat A, Duaibis R, et al. Frequency estimates and risk factors for postoperative morbidity after third molar removal: a prospective cohort study. *J Oral Maxillofac Surg* 2008 Nov;66(11):2276-2283. [doi: [10.1016/j.joms.2008.06.047](https://doi.org/10.1016/j.joms.2008.06.047)] [Medline: [18940492](https://pubmed.ncbi.nlm.nih.gov/18940492/)]
25. MacGregor A. Aetiology of dry socket: A clinical investigation. *British Journal of Oral Surgery* 1968 Jan;6(1):49-58. [doi: [10.1016/S0007-117X\(68\)80026-5](https://doi.org/10.1016/S0007-117X(68)80026-5)]
26. Hedström L. Dry socket--an evaluation of 163 cases of the surgical removal of mandibular wisdom teeth. *Tandlaktidningen* 1987 Mar;79(6):298-303. [Medline: [3473729](https://pubmed.ncbi.nlm.nih.gov/3473729/)]
27. Blum I. Contemporary views on dry socket (alveolar osteitis): a clinical appraisal of standardization, aetiopathogenesis and management: a critical review. *Int J Oral Maxillofac Surg* 2002 Jun;31(3):309-317. [doi: [10.1054/ijom.2002.0263](https://doi.org/10.1054/ijom.2002.0263)] [Medline: [12190139](https://pubmed.ncbi.nlm.nih.gov/12190139/)]
28. Bortoluzzi M, Manfro R, De Déa BE, Dutra T. Incidence of dry socket, alveolar infection, and postoperative pain following the extraction of erupted teeth. *J Contemp Dent Pract* 2010 Jan 01;11(1):E033-E040. [Medline: [20098964](https://pubmed.ncbi.nlm.nih.gov/20098964/)]
29. Zozus M, Hammond W, Richesson R. NIH Collaboratory. 2014. Phenotypes, Data Standards, and Data Quality Core URL: https://www.nihcollaboratory.org/Documents/Hammond_Zozus_2014-0816-PDSDQCore_update.pdf [accessed 2019-03-18] [WebCite Cache ID 76y62LphV]
30. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. *Int J Environ Res Public Health* 2014 May 14;11(5):5170-5207 [FREE Full text] [doi: [10.3390/ijerph110505170](https://doi.org/10.3390/ijerph110505170)] [Medline: [24830450](https://pubmed.ncbi.nlm.nih.gov/24830450/)]
31. Faulconer E, de Lusignan S. An eight-step method for assessing diagnostic data quality in practice: chronic obstructive pulmonary disease as an exemplar. *Inform Prim Care* 2004;12(4):243-254 [FREE Full text] [Medline: [15808026](https://pubmed.ncbi.nlm.nih.gov/15808026/)]
32. Booth H, Prevost A, Gulliford M. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. *Pharmacoepidemiol Drug Saf* 2013 Dec;22(12):1357-1361 [FREE Full text] [doi: [10.1002/pds.3537](https://doi.org/10.1002/pds.3537)] [Medline: [24243711](https://pubmed.ncbi.nlm.nih.gov/24243711/)]

33. McVeigh K, Newton-Dame R, Chan P, Thorpe L, Schreiberstein L, Tatem K, et al. Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data. *EGEMS (Wash DC)* 2016;4(1):1267 [FREE Full text] [doi: [10.13063/2327-9214.1267](https://doi.org/10.13063/2327-9214.1267)] [Medline: [28154837](https://pubmed.ncbi.nlm.nih.gov/28154837/)]
34. Perlman S, McVeigh K, Thorpe L, Jacobson L, Greene C, Gwynn R. Innovations in Population Health Surveillance: Using Electronic Health Records for Chronic Disease Surveillance. *Am J Public Health* 2017 Dec;107(6):853-857. [doi: [10.2105/AJPH.2017.303813](https://doi.org/10.2105/AJPH.2017.303813)] [Medline: [28426302](https://pubmed.ncbi.nlm.nih.gov/28426302/)]
35. Thorpe L, McVeigh K, Perlman S, Chan P, Bartley K, Schreiberstein L, et al. Monitoring Prevalence, Treatment, and Control of Metabolic Conditions in New York City Adults Using 2013 Primary Care Electronic Health Records: A Surveillance Validation Study. *EGEMS (Wash DC)* 2016;4(1):1266 [FREE Full text] [doi: [10.13063/2327-9214.1266](https://doi.org/10.13063/2327-9214.1266)] [Medline: [28154836](https://pubmed.ncbi.nlm.nih.gov/28154836/)]
36. Akinbami B, Godspower T. Dry socket: incidence, clinical features, and predisposing factors. *Int J Dent* 2014;2014:796102 [FREE Full text] [doi: [10.1155/2014/796102](https://doi.org/10.1155/2014/796102)] [Medline: [24987419](https://pubmed.ncbi.nlm.nih.gov/24987419/)]
37. Kolokythas A, Olech E, Miloro M. Alveolar osteitis: a comprehensive review of concepts and controversies. *Int J Dent* 2010;2010:249073 [FREE Full text] [doi: [10.1155/2010/249073](https://doi.org/10.1155/2010/249073)] [Medline: [20652078](https://pubmed.ncbi.nlm.nih.gov/20652078/)]
38. Abu Younis MH, Abu Hantash RO. Dry socket: frequency, clinical picture, and risk factors in a palestinian dental teaching center. *Open Dent J* 2011 Feb 07;5:7-12 [FREE Full text] [doi: [10.2174/1874210601105010007](https://doi.org/10.2174/1874210601105010007)] [Medline: [21559187](https://pubmed.ncbi.nlm.nih.gov/21559187/)]
39. Sweet J, Butler D. Predisposing and operative factors: effect on the incidence of localized osteitis in mandibular third-molar surgery. *Oral Surg Oral Med Oral Pathol* 1978 Aug;46(2):206-215. [Medline: [280828](https://pubmed.ncbi.nlm.nih.gov/280828/)]
40. Steckler A, McLeroy K. The importance of external validity. *Am J Public Health* 2008 Jan;98(1):9-10. [doi: [10.2105/AJPH.2007.126847](https://doi.org/10.2105/AJPH.2007.126847)] [Medline: [18048772](https://pubmed.ncbi.nlm.nih.gov/18048772/)]
41. Osborn J, Stoltenberg J, Newell K, Osborn S. Adequacy of dental records in clinical practice: a survey of dentists. *J Dent Hyg* 2000;74(4):297-306. [Medline: [11314481](https://pubmed.ncbi.nlm.nih.gov/11314481/)]
42. Helminen S, Vehkalahti M, Murtomaa H. Dentists' perception of their treatment practices versus documented evidence. *Int Dent J* 2002 Apr;52(2):71-74. [Medline: [12013253](https://pubmed.ncbi.nlm.nih.gov/12013253/)]
43. Majeed A, Car J, Sheikh A. Accuracy and completeness of electronic patient records in primary care. *Fam Pract* 2008 Aug;25(4):213-214. [doi: [10.1093/fampra/cmn047](https://doi.org/10.1093/fampra/cmn047)] [Medline: [18694896](https://pubmed.ncbi.nlm.nih.gov/18694896/)]
44. Oksuzyan A, Demirchyan A, Thompson M. Evaluation of Medical Records Completeness in the Adult Cardiology Clinic at Nork Marash Medical Center. Yerevan, Armenia: American University of Armenia; 2003 Aug. URL: http://aua.am/chsr/UserFiles/File/Anahit/Evaluation_of_SEF_Completeness_Eng.PDF [accessed 2019-03-16] [WebCite Cache ID 74Xzr2c4i]
45. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 2003 May 17;326(7398):1070 [FREE Full text] [doi: [10.1136/bmj.326.7398.1070](https://doi.org/10.1136/bmj.326.7398.1070)] [Medline: [12750210](https://pubmed.ncbi.nlm.nih.gov/12750210/)]
46. de Mul M, Berg M. Completeness of medical records in emergency trauma care and an IT-based strategy for improvement. *Med Inform Internet Med* 2007 Jun;32(2):157-167. [doi: [10.1080/09670260701231284](https://doi.org/10.1080/09670260701231284)] [Medline: [17541865](https://pubmed.ncbi.nlm.nih.gov/17541865/)]
47. Tola K, Abebe H, Gebremariam Y, Jikamo B. Improving Completeness of Inpatient Medical Records in Menelik II Referral Hospital, Addis Ababa, Ethiopia. *Advances in Public Health* 2017;2017:1-5. [doi: [10.1155/2017/8389414](https://doi.org/10.1155/2017/8389414)]
48. University of California. 2008. Legal Medical Record Standards URL: <http://policy.ucop.edu/doc/1100168/LegalMedicalRecord> [accessed 2019-03-16] [WebCite Cache ID 74Xu6Rjuo]
49. Waleed P, Baba F, Alsulami S, Tarakji B. Importance of dental records in forensic dental identification. *Acta Inform Med* 2015 Feb;23(1):49-52 [FREE Full text] [doi: [10.5455/aim.2015.23.49-52](https://doi.org/10.5455/aim.2015.23.49-52)] [Medline: [25870492](https://pubmed.ncbi.nlm.nih.gov/25870492/)]
50. Thierer TE, Delander KA. Improving Documentation, Compliance, and Approvals in an Electronic Dental Record at a U.S. Dental School. *J Dent Educ* 2017 Apr;81(4):442-449 [FREE Full text] [doi: [10.21815/JDE.016.004](https://doi.org/10.21815/JDE.016.004)] [Medline: [28365609](https://pubmed.ncbi.nlm.nih.gov/28365609/)]
51. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. *Diabetologia* 2018 Dec;61(6):1241-1248. [doi: [10.1007/s00125-017-4518-6](https://doi.org/10.1007/s00125-017-4518-6)] [Medline: [29247363](https://pubmed.ncbi.nlm.nih.gov/29247363/)]
52. Casey J, Schwartz B, Stewart W, Adler N. Using Electronic Health Records for Population Health Research: A Review of Methods and Applications. *Annu Rev Public Health* 2016;37:61-81. [doi: [10.1146/annurev-publhealth-032315-021353](https://doi.org/10.1146/annurev-publhealth-032315-021353)] [Medline: [26667605](https://pubmed.ncbi.nlm.nih.gov/26667605/)]
53. Yoon S, Odlum M, Lee Y, Choi T, Kronish I, Davidson K, et al. Applying Deep Learning to Understand Predictors of Tooth Mobility Among Urban Latinos. *Stud Health Technol Inform* 2018;251:241-244 [FREE Full text] [Medline: [29968648](https://pubmed.ncbi.nlm.nih.gov/29968648/)]
54. Moskovitch R, Polubriaginof F, Weiss A, Ryan P, Tatonetti N. Procedure prediction from symbolic Electronic Health Records via time intervals analytics. *J Biomed Inform* 2017 Nov;75:70-82 [FREE Full text] [doi: [10.1016/j.jbi.2017.07.018](https://doi.org/10.1016/j.jbi.2017.07.018)] [Medline: [28823923](https://pubmed.ncbi.nlm.nih.gov/28823923/)]
55. Chan P, Zhao Y, Lim S, Perlman S, McVeigh K. Using Calibration to Reduce Measurement Error in Prevalence Estimates Based on Electronic Health Records. *Prev Chronic Dis* 2018 Dec 13;15:E155 [FREE Full text] [doi: [10.5888/pcd15.180371](https://doi.org/10.5888/pcd15.180371)] [Medline: [30576279](https://pubmed.ncbi.nlm.nih.gov/30576279/)]
56. Hripcsak G, Duke J, Shah N, Reich C, Huser V, Schuemie M, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform* 2015;216:574-578 [FREE Full text] [Medline: [26262116](https://pubmed.ncbi.nlm.nih.gov/26262116/)]
57. Boeckhout M, Zielhuis G, Bredenoord A. The FAIR guiding principles for data stewardship: fair enough? *Eur J Hum Genet* 2018 Dec;26(7):931-936 [FREE Full text] [doi: [10.1038/s41431-018-0160-0](https://doi.org/10.1038/s41431-018-0160-0)] [Medline: [29777206](https://pubmed.ncbi.nlm.nih.gov/29777206/)]

58. Cervero R, Gaines J. Accreditation Council for Continuing Medical Education. 2014. Effectiveness of Continuing Medical Education: Updated Synthesis of Systematic Reviews URL: http://www.accme.org/sites/default/files/2014_Effectiveness_of_Continuing_Medical_Education_Cervero_and_Gaines_0.pdf [accessed 2019-03-16] [WebCite Cache ID 74Y03uqEp]
59. Bates D, Kuperman G, Wang S, Gandhi T, Kittler S, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10(6):523-530 [FREE Full text] [doi: [10.1197/jamia.M1370](https://doi.org/10.1197/jamia.M1370)] [Medline: [12925543](https://pubmed.ncbi.nlm.nih.gov/12925543/)]

Abbreviations

AEGD: advanced education in general dentistry
EDR: electronic dental record
EMR: electronic medical record
GPR: general practice residency
OMFS: Department of Oral and Maxillofacial Surgery

Edited by C Lovis; submitted 17.12.18; peer-reviewed by F Shaya, R Moskovitch, M El Tantawi; comments to author 28.12.18; revised version received 11.02.19; accepted 13.03.19; published 21.03.19.

Please cite as:

Levitin SA, Grbic JT, Finkelstein J

Completeness of Electronic Dental Records in a Student Clinic: Retrospective Analysis

JMIR Med Inform 2019;7(1):e13008

URL: <http://medinform.jmir.org/2019/1/e13008/>

doi: [10.2196/13008](https://doi.org/10.2196/13008)

PMID: [30896435](https://pubmed.ncbi.nlm.nih.gov/30896435/)

©Seth Aaron Levitin, John T Grbic, Joseph Finkelstein. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 21.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

OpenEHR and General Data Protection Regulation: Evaluation of Principles and Requirements

Duarte Gonçalves-Ferreira^{1,2*}, MS; Mariana Sousa^{1,2*}, MSc; Gustavo M Bacelar-Silva¹, MSc, MD, MBA; Samuel Frade¹, MSc; Luís Filipe Antunes², PhD; Thomas Beale³, BE, BSc; Ricardo Cruz-Correia¹, PhD

¹Center for Health Technology and Services Research, Porto, Portugal

²Healthy Systems, Porto, Portugal

³Ars Semantica, London, United Kingdom

*these authors contributed equally

Corresponding Author:

Duarte Gonçalves-Ferreira, MS

Center for Health Technology and Services Research

R Dr Plácido da Costa

Porto, 4200-450

Portugal

Phone: 351 917681423

Email: dferreira@fe.up.pt

Abstract

Background: Concerns about privacy and personal data protection resulted in reforms of the existing legislation in the European Union (EU). The General Data Protection Regulation (GDPR) aims to reform the existing directive on the topic of personal data protection of EU citizens with a strong emphasis on more control of the citizens over their data and in the establishment of rules for the processing of personal data. OpenEHR is a standard that embodies many principles of interoperable and secure software for electronic health records (EHRs) and has been advocated as the best approach for the development of hospital information systems.

Objective: This study aimed to understand to what extent the openEHR standard can help in the compliance of EHR systems to the GDPR requirements.

Methods: A list of requirements for an EHR to support GDPR compliance and also a list of the openEHR design principles were made. The requirements were categorized and compared with the principles by experts on openEHR and GDPR.

Results: A total of 50 GDPR requirements and 8 openEHR design principles were identified. The openEHR principles conformed to 30% (15/50) of GDPR requirements. All the openEHR principles were aligned with GDPR requirements.

Conclusions: This study showed that the openEHR principles conform well to GDPR, underlining the common wisdom that truly realizing security and privacy requires it to be built in from the start. By using an openEHR-based EHR, the institutions are closer to becoming compliant with GDPR while safeguarding the medical data.

(*JMIR Med Inform* 2019;7(1):e9845) doi:[10.2196/medinform.9845](https://doi.org/10.2196/medinform.9845)

KEYWORDS

health information interoperability; electronic health record; data protection; GDPR; openEHR

Introduction

The computer-based patient record has been considered an essential technology for health care in the last 25 years [1] even though their cost-effectiveness still needs more research to be fully assessed [2]. Currently, health care activities strongly rely on collected patient data and are feeding big data-driven health care projects that, among other aims, seek reliable predictors of health outcomes [3]. Health care professionals deal with a

great volume of data, as their activities are heavily dependent on information accessed, as well as the way it is processed, managed, and made available.

Information technology (IT) development has enabled health care institutions to improve the collection and processing of health data, raising new concerns regarding the sensitivity of the information processed by information systems (ISs), namely, the risks concerning patient data protection and privacy.

Although easy access to information is crucial to routine clinical practice, privacy, and security of medical information, it cannot be neglected, considering the consequences the misuse of medical information can present to the patient's personal life. The Health Insurance Portability and Accountability Act privacy and security rules clearly emphasize the need of privacy of health information of the patient while allowing for sharing among different agencies [4].

One example of patient data misuse is the use of medical records for research without consent. This misuse is unfortunately widespread in the institutions that we have contact with. Often, this is a concern after the patient data are already accessed and just before the research is sent for publishing. This attracts little scrutiny compared with, for instance, biospecimen research, where concerns about genomic privacy prompted recent US federal proposals to mandate consent [5]. Moreover, cybersecurity threats against health care organizations are rising in numbers and severity [6,7]. New threats include hacktivism and organized crime, targeting individual identified data stored in hospitals and offices. New technologies and policies are needed to address the risks [8].

Health care data standards such as HL7 v2, HL7 FHIR [9], DICOM [10], or openEHR are central in the quality of the implemented patient records and have also tried to address security issues. Nevertheless, the pressure on health care to comply to new data protection rules is rising, so these standards must also be revisited taking this into consideration.

General Data Protection Regulation

The European Union (EU) General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years [11]. The GDPR replaces the Data Protection Directive 95/46/EC in an increasingly data-driven world that is vastly different from the time in which the 1995 directive was established. The GDPR was designed to harmonize data privacy laws across Europe to protect and empower all EU citizens' data privacy and to reshape the way organizations across the region approach data privacy. Although the key principles of data privacy still hold true to the previous directive, many changes have been proposed to the regulatory policies. The GDPR was adopted on 27 April 2016. It becomes enforceable from 25 May 2018, after a 2-year transition period at which time those organizations in noncompliance will face heavy fines. So, it is of utmost importance for health care institutions to acknowledge the regulation's requirements, analyze what is imposed in the obligations, and verify the compliance of the organization and its ISs, as well as, when necessary, to define strategies to adopt necessary measures.

Regarding principles relating to processing of personal data, the GDPR includes lawfulness and transparency toward the data subject; there must be a clear purpose for data collection and limitation regarding the further processing of data other than archiving; data processing should be adequate, relevant, and limited to what is necessary, fulfilling the principle of data minimization; personal data should be accurate and, when necessary, kept up-to-date (accuracy); personal data should be kept in a form, which permits identification of data subjects for no longer than is necessary (storage limitation); personal data

should be processed in a manner that ensures its appropriate security (integrity and confidentiality); and the controller shall be responsible for and be able to demonstrate compliance with the regulation (accountability).

Regarding the rights of the data subjects, the GDPR defines the rights that controllers must make possible, such as the rights to be informed, of access by the data subject, to rectification, to erasure (commonly known as the *right to be forgotten*), to portability, to limitation to processing, to restriction of processing, to not be subject to automated individual decision making, including profiling, and to withdraw consent for processing of data. The controllers are obliged to implement technical and organizational measures that can answer to the data subject's request, regarding their rights, but also at a security level.

Organizations must incorporate concepts such as privacy by design and by default in the development of their systems, to comply with GDPR requirements related to the protection of the personal data they process. Even though these requirements are seen as a restriction for medical research, there are pointers in the literature to the standardization of the data in Europe and uniformization of a digital single market together with the GDPR, which will facilitate medical research when the research is considered in the public interest [12].

OpenEHR

OpenEHR presents a set of principles for an interoperable EHR systems architecture based on a multilevel, single source modeling approach. OpenEHR's specifications are published by the openEHR Foundation, an entity responsible for the development of the specifications and the availability of specific tools enabling the standard's use. One of the main goals of openEHR is to enable the development of EHR systems to be able to communicate with each other, without loss of meaning, thus achieving semantic interoperability.

Modeling in openEHR relies on a 2-level scheme that separates the content from the form in which it is defined. The openEHR Architecture Overview states that, under the 2-level approach, the first level is a stable reference information model, which defines basic concepts for logical data representation, which also act as primitives for the second layer of models. These primitives include data types, structures, and the connections between them controlling how they can be assembled to create clinical content definitions in the second level. The clinical content definitions consist of data points, and groups are defined in the form of constraint structures, known as *archetypes*, on the first layer primitives. Archetypes can be used to create other more complex archetypes and also *templates* which are representations of datasets for specific domain use cases. Thus, an openEHR Archetype is the model (or pattern) for the capture of clinical information—machine-readable specification of how to store patient data using the openEHR Reference Model whereas archetypes describe complete domain-level data structures such as *diagnosis* or *test result*, and a template provides the means for grouping archetype-defined data points for particular business purposes.

With this 2-level approach, the clinical content is structured *outside the software*, allowing EHR systems to be more flexible, as the modifications concerning the clinical knowledge are realized solely by the modifications of the archetypes, without compromising the integrity of the software or data of an EHR repository which is based on the Reference Model [13].

OpenEHR offers a new paradigm of systems modeling, relying on a very stable model at the software level and a very flexible modeling that reflects the evolution of knowledge at the domain level. There are tools to help the modeling process such as the Clinical Knowledge Manager which is a Web-based repository that contains archetypes and templates developed by an international group of specialists. This platform supports collaboration open to everyone (specially clinicians, IT professionals, and software engineers), where participants can author, review, translate, and maintain archetypes and templates.

Aim

Given the nature of openEHR as a standard being used to build EHR systems, it is important to understand to what extent the openEHR principles address the requirements mandatory to GDPR. This research aimed to study if and how openEHR addresses the GDPR requirements.

Methods

The study was performed in 3 steps: (1) identify the requirements for a health information system (HIS) compliant with GDPR; (2) identify the openEHR security principles regarding the functionalities of an HIS; and (3) determine the correspondence of the openEHR principles to the GDPR requirements.

General Data Protection Regulation Requirements

The list of the GDPR requirements was created by reading the legislation by specialists (authors of this paper: LFA and MS). The list of requirements was built with a strong input on the global description of the system, focusing on the identification of the GDPR goals and the later translation to system functionalities. The requirements were described using the Institute of Electrical and Electronics Engineers Guide for Developing System Requirements Specifications [14].

A search was conducted on the PubMed database for papers related to the GDPR using the GDPR keyword. The search returned a list of 29 papers from which the ones without an abstract as well as the ones that did not relate specifically to the subject in hand were removed. We were left with 5 papers that were reviewed to obtain higher level groups for the GDPR requirements.

OpenEHR Principles

The list of openEHR architectural features relevant to GDPR was compiled from the openEHR Architecture Overview [13] by specialists (authors of this paper: GB and SF), aiming to identify its main principles in view of the functionalities of an openEHR-based system. A description of each principle was agreed upon by the specialists.

We identified and listed the openEHR features, with a strong focus on the functionalities of a system rather than the implications of the architecture.

Matching General Data Protection Regulation Requirements With OpenEHR Features

Each feature can match more than one requirement, and a requirement can be matched by more than one feature. To be considered as a match, the openEHR features should meet the GDPR requirements in a straightforward way by the simple implementation of its architecture.

Results

The results section presents the (1) list of the GDPR requirements, (2) the openEHR GDPR-related features, (3) a table that matches the requirements with the features, and (4) a list of requirements not met by openEHR GDPR-related features.

General Data Protection Regulation Requirements

The article review on the GDPR in PubMed left us with a list of 5 articles that were relevant for the subject of GDPR in health care. Moreover, 2 of the 5 articles focus on data sharing [15,16] and set Consent, Privacy, Security Measures, Adequacy of use, and Oversight as a high-level grouping of GDPR concerns in health care data sharing. Furthermore, another 2 articles were related to the concerns of the GDPR in research [17] and in the area of radiology [10]. Although the first focus was on a review of the GDPR for medical research, it does not provide a usable division of the requirements, focusing instead on the changes GDPR brings to researchers; the second focus was on term definitions such as portability of health care data, personal data breaches, anonymization, pseudonymization, and encryption, in which requirements affect the lawful processing of data for research. The final article describes a system that focuses on the audit and traceability helping institutions fulfill the need that GDPR imposes in the institution to know who, when, and what is done with their data [18]. On top of these articles, we analyzed an article by Mense and Blobel [9] where the authors analyze the GDPR requirements, extract some key factors from the legislation, and match them against multiple HL7 standards including CDA, FHIR, and HL7 v2. They synthesize the GDPR legislation into 7 key factors that we enumerate next:

1. Data protection by design and by default.
2. Data portability.
3. Right to be forgotten—notification requirement.
4. Unambiguous consent.
5. Privacy notices.
6. Right to Access and Records of processing activities.
7. Explicit and formally represented policies.

In our analyses of the GDPR legislation [11], we identified a total of 50 requirements grouped into 7 groups. The requirements were aggregated into the following groups:

1. Limitations to data processing, which include requirements directly related with data processing limitation for the institution.

2. Data quality and accountability, which include requirements related with integrity, accuracy, and audit.
3. Consent by data subject, requirements related with consent and authorization.
4. Empowerment of data subject, requirements that increase the rights of the data subjects on their data.
5. Data breaches, requirements directly related with data breaches and how to proceed in case of a data breach.
6. Data portability and interoperability, requirements related to authorized data sharing.
7. Privacy control and impact assessment, requirements related with Privacy Impact Assessment (PIA) and privacy by default and by design.

The complete list of the requirements is defined in [Textbox 1](#).

OpenEHR Design Principles

The following 8 features were identified in openEHR as being relevant to GDPR:

Feature 1: 2-Level Modeling

2-level modeling promotes the separation of the reference model (stable information model that defines the logic structure of the EHR and the demographic data) from the content model (the definition of clinical domain as archetypes and templates modeled by clinical professionals. Essentially, archetypes and templates are datasets external to any system's software). The Reference Model is implemented at the software level, whereas the Domain Model is set through the archetype and template modeling. This results in the separation and independence of software structure from its content, enabling flexible, interoperable, and scalable health systems. Fundamentally, all openEHR systems support the same data structure and remain able to communicate, regardless of how many changes are made to domain information definitions.

Feature 2: Separation of Clinical and Demographic Information

One of the openEHR design principles is to enable the complete separation of the EHR from identifiable demographic information via separated repositories with flexible referencing. In case of a data breach of the EHR repository, it allows the identity of the data subject to be preserved, unless the demographic repository is also breached. This principle strengthens the data subject's anonymization regarding the information in their EHR, as it is used as an instance in the EHR, called *Party_Self*, to make a reference to the data subject. This information works as an optional external reference, such that the EHR can be set to provide 3 levels of separation. The external identifier is determined in the instance *Party_Self* by

- Nowhere in the EHR (every *Party_Self* instance is left empty). This is the safest way, and it means that the connection between the EHR and the patient needs to be

made outside the EHR, by connecting the EHR identifier (EHR.ehr_id) and the subject's identifier.

- Only once in the *EHR_STATUS* object (subject's attribute) and nowhere else. It is a very safe measure if the *EHR_STATUS* object is protected in any way.
- In any *Party_Self* instance, this is a reasonable solution in a safe environment appropriate to copy parts of the record on demand.

Feature 3: Service Model

openEHR service model [19,20] specifies a formal, abstract definition of interfaces to be implemented in an openEHR system. Implementations can follow these abstract definitions to allow interoperability between various implementations. The service model currently consists of the EHR Service, the Query Service, and the Definitions Service. The EHR Service allows the consultation of data made available by the EHR Application Programming Interface (API). The detail level of the consulted data may vary, allowing the access to more complex records, such as changes in versions, or it may allow the search of simpler elements, such as single clinical data items, patient identifications, etc. The Definitions Service allows access to an archetype repository, acting as an important tool for access to important information by medical professionals (eg, if they need an archetype that is not on his/her local repository for certain medical treatment). The service model thus takes on an important role in controlling the availability of data, as well as the possibility of consultation, allowing the definition of safe and intuitive views.

Features 4 and 5: Version Control—Versioning and Digital Signature

Important openEHR features related to GDPR relate to data integrity support. The EHR or demographic repository is managed using *Versioned Objects*. Versioned Objects are used to contain the versions of a *Composition* or *Party* structure, which in turn contain fine-grained clinical and demographic data, respectively. The set of changes to items in any update to the system is called a *Contribution* (more commonly known as a *change set*). Each change set works as a transaction, ensuring the consistency and integrity of the data repository. Changes made by users (creating new records, deleting records, modifying records, and transferring records) are not performed at the item/record level but at the level of the repository as a whole. This means that no version is deleted or modified; all the changes are physically implemented as new versions in the repository. This principle ensures indelibility (no information is deleted). Version control includes the possibility of each version having a digital signature, created as a primary-key encrypted of a hash of a representation approved of the compromised version. In a versioned system, the digital signature acts as a verification of integrity, a measure of authentication and also as a measure of nonrepudiation.

Textbox 1. List of General Data Protection Regulation requirements.

1. Limitations to data processing
 - Purpose limitation: The system shall admit the definition of a purpose for the limitation of processing.
 - Data minimization: The system must allow the definition of the minimum of data fields required for processing.
 - Period of storage limitation: The system must allow the definition of deadlines for the processing of specific personal data, in order with the purpose of processing.
 - Method storage limitation: The system must allow the storage of data in a way that only identifies the data subjects during the necessary time relative to the purpose.
 - Limitation of processing of personal data: The system must be able to limit the processing of personal data according to the consent given by the data subject.
2. Data quality and accountability
 - Accuracy: The system must allow the update of the personal data whenever necessary.
 - Integrity and confidentiality: The system must support the adoption of technical and organizational measures that ensure the security of processing, namely, the protection against unauthorized processing or against the loss, destruction, or accidental harm of personal data.
 - Accountability: The system must allow the demonstration of compliance with the data processing principles.
 - Statement of accountability: The system must support the demonstration of compliance with codes of conduct and certified procedures.
 - Conditions of processing: The system must record data describing the legal context that allows the processing of data.
 - Record of processing: The system shall be able to keep an up-to-date and accurate record of all the processing activities and must allow the record of processing to be written in an electronic format.
 - Availability of records of processing: The system must allow access to consult its records of processing.
 - Location of data: The system must be able to identify and locate a subject's data that must be limited inside the system.
3. Consent by data subject
 - Explicit consent: The system must be able to record and show the consent of the data subjects for the collection of personal data and the purpose for collecting.
 - Management of consent: The system must allow changes to the consent by the data subject.
 - Record of consent: The system must be able to keep a record of consent or consents to distinguish it from other content.
 - Withdrawal of consent: The system must ensure the ability to withdraw consent (opt-out) in an easy and clear way, using the same means in which the consent was obtained.
 - Features of the consent: The system must ensure that the consent provision by a subject is active, not obtained through silence, inactivity, or prechecked boxes, and that it is confirmed in words.
 - Lawfulness of processing after withdrawal of consent: The system shall be able to ensure the lawfulness of data processing after the withdrawal of consent.
 - Objection to processing: The system must support the cessation of processing in response to a request by the data subject.
4. Data subject empowerment
 - Information provided to data subject: The system must inform the data subject about the conditions and rules relating to data processing and privacy.
 - Means to provide information to data subjects: The system must have a means of providing such information.
 - Verification of the identity of the data subject: The system must allow the verification of the identity of data subjects upon the request.
 - Data subject request: The system must support the receipt of data subject's request.
 - Response to request: The system must enable the solicitation to the data subject's request by the same means the request was made.
 - Data subject access: The system must provide a copy of the data subject's personal data at processing on request.
 - Data subject request action: The systems must support the means for the data subject to request access to the subject's data.
 - Information accessed by the data subject—The system must enable access to the data subject's information and actions such as the following:
 - Purpose of processing.
 - Categories of personal data held by the system.
 - The recipients or categories of recipients to whom personal data have been or will be disclosed.

- The period for which personal data will be stored.
 - The existence of the right to request from the controller rectification or erasure or restriction of data processing.
 - Right to lodge a complaint with a supervisory authority.
 - Available information as to the source of data collection, if personal data were not collected from the data subject.
 - Existence of automated decision making, including profiling.
- Response to data subject request: The system must allow the response to the data subject's request in a commonly used means.
 - Data subject direct access: The system must provide a secure method for the direct access of the data subject to their personal data.
 - Personal data rectification: The system must allow the rectification of inaccurate personal data by the data subject.
 - Personal data erasure: The system must allow the erasure of personal data when consent is removed or when the purpose on which the data were gathered is no longer valid.
 - Legitimate interest: The system must be able to demonstrate to the data subject the legitimate interest for the processing, including retrieval, modification, and sharing.
 - Confirmation of data processing: The system must be able to confirm the processing of the data subject's personal data in each case requested by the subject.
5. Data breaches
- Records of data breaches: The system must keep a record of data breaches that were detected.
 - Records of data breaches nature: The system must register information regarding the nature of data breach.
 - Data breach description: The system must keep a record of information regarding the nature of data breach in a format subject to be sent to the supervisory authority.
 - Data breach notification deadline: The system must enable the notification procedure of the supervisory authority in 72 hours.
 - Data breaches notification procedures: The system must support the development of procedures for the report of internal breaches.
6. Data portability and interoperability
- Portability of personal data: The system must allow the portability of personal data in a structured, common, automatic format.
 - Portability of personal data between controllers: The system must be able to transfer personal data to another controller.
 - Interoperability of systems and formats: The system must enable interoperability for the transfer and portability of personal data.
 - Communication between institutions: The system must allow the communication between institutions involved in the processing of the same personal data.
 - Cross-border data transfers: The system must allow the transfer of personal data to other countries.
 - Cross-border data transfers guarantee: The system must enable the recording of the proper measures presented by the third country or international organization that allows the transfer of personal data.
7. Privacy control and impact assessment
- Privacy by design: The system must allow the pseudonymization and encryption of data and must be able to apply data minimization measures, storing only minimal needed data.
 - Privacy by default: The system must ensure the processing of personal data relevant to the purpose and it must ensure that no personal data are made available without human intervention.
 - Access control measures: The system must make data unintelligible in case of unauthorized access.
 - Data Protection Impact Assessment (DPIA) records preservation: The system must allow the preservation of DPIA.
 - DPIA consultation: The system must allow the consultation of the DPIA when the controller requires it.

Features 6 and 7: Access Control—Access Control List and Configurations

openEHR access control is set through the object named *EHR_ACCESS*. This object works as a gate to all access information, being that any decision regarding access information should be based on the policies and rules established in it. OpenEHR's EHR allows the definition of an access control

list, indicating the identified individuals and their categories. The definition of the access control list should consider relevance of the user's identity access either in terms of time or duration of access. When creating an EHR, openEHR allows the definition of a gatekeeper responsible for the access control configurations. The gatekeeper becomes an identity recognized in the EHR, usually being the own patient (in case of mentally competent adults) or a relative or legal tutor (in case of being

a child or mentally incapable). The gatekeeper sets who can make changes in the access control list, being that all changes are kept in the audit trail. These features could make use of recent developments in security technologies referred to in the literature [21,22], although they still need field validation in the health care area.

Feature 8: Audit Trailing

All changes that are made, at all levels, in the EHR are recorded in the audit trail, with data related to the identity of the user, timestamp, purpose (of the alterations performed), digital signature, and relevant version information.

Matching General Data Protection Regulation Requirements With OpenEHR Principles

Table 1 presents the existing matches between the GDPR requirements and the openEHR principles.

The results obtained showed that openEHR GDPR-related features satisfied at least 1 identified GDPR requirement.

The GDPR requirement Method Storage Limitation (1.4), listed in [Textbox 1](#), is fulfilled by the Separation of EHR and demographic information, allowing a separate storage of demographic and clinical data. The identity of the data subject is automatically preserved when the clinical and demographic information are separated. In that way, although the clinical data are stored for treatment, the demographic data are only connected to the EHR through an external identifier, allowing the identification of the data subject only during the necessary period of the purpose of processing (typically only on a device used by an authorized health professional during a health care process).

Integrity and Confidentiality (2.2), listed in [Textbox 1](#), is a complex requirement in the GDPR and is thus fulfilled by a group of openEHR architectural features. Data versioning allows the creation of new EHR versions, ensuring indelibility, and is thus an important measure against the loss, destruction, or accidental corruption of EHR data, guaranteeing trustworthy and reliable information at all moments of processing. The digital signing of data ensures the authentication, nonrepudiation, and integrity of the EHR, acting as an important security and integrity measure of the personal data and its processing. The access control rules are included in the openEHR architecture by design, and it ensures the confidentiality of the patient's data by limiting the occurrence of an unauthorized or illicit processing, because of the definition of who is authorized to access the data. There is also a component that identifies who can access or change the access control rules, it sets the individuals who can change the configurations of the access list, contributing to legitimate and justified accesses and ensuring the integrity of processing. On top of these preventive features, the openEHR services architecture [23] includes a System Log service that records access actions, as well as identity of the users, date and time, and justification of the action, ensuring the integrity of the data and postaccess analyses. (This service is not defined by openEHR but assumed to be an implementation of, eg, IHE ATNA.) These design features ensure the security of processing,

correctly identifying when, who, what, and how data were accessed, and allowing a postaccess audit.

Accountability (2.3), Record of processing (2.6), and Availability of records of processing (2.7), listed in [Textbox 1](#), are fulfilled by the audit trail feature, which allows the system to keep a record of all information related to the processing of data. This way, it is possible to demonstrate compliance with the principles and obligation of the requirements, as it is possible to see information related to access and actions taken in the EHR. Due to its traceability, it allows the creation of a record of personal data processing that can become available to the data authorities.

The requirement Verification of the identity of the data subjects (4.3), listed in [Textbox 1](#), is fulfilled by the separation of demographic and clinical data that ensures the separation of identifiable information from the clinical data. The EHR has an identifier associated to a single patient, ensuring the identification of the data subject's identity if necessary.

Data subject access (4.6) and Data subject direct access (4.10), listed in [Textbox 1](#), are two of the requirements that we identified for the GDPR and can be fulfilled by the Access control feature of the openEHR. The data subject can be granted access using the access control list and can manage this list through the Access control configurations. Data subject direct access is fulfilled by an extra feature, the Service Model feature that allows the creation of direct access reference to the subject's data by the subject.

Confirmation of data processing (4.14), listed in [Textbox 1](#), requirement is fulfilled by the audit trailing feature of the openEHR principles. In the audit trail, systems can store every action that is done on the subject's data. Also, it is possible to check what changes were done to the data through the different versions created in each action.

Data portability and interoperability are, by design, a key focus of the openEHR architecture and 2-level modeling, so the GDPR requirements 6.1, 6.2, 6.3, and 6.5 can be matched by these features. The interoperability point of the requirement 6.3 is also fulfilled with the integration of another of the openEHR principles, the Service Model that allows the creation of different interfaces, in different systems around the institution, using the same data. By supporting different views that allow the consultation of the same EHR, the record maintains its integrity and structure, ensuring interoperability.

Privacy by design and by default were identified as requirements 7.1 and 7.2 in [Textbox 1](#). The separation of demographic and clinical data by design improves the protection of the data subject identification by separating the EHR from the identifiable demographic information, only relating them by an external identifier. By default, during health care, only the EHR data should be considered. Privacy by default is also matched by the Access control list and its configuration that endures the access and availability of information and ensures that the data are only processed and accessed for the purpose settled, by those authorized to do it, safeguarding the patient's data privacy.

Table 1. List of the 17 General Data Protection Regulation (GDPR) requirements that are met by openEHR principles.

GDPR requirements	openEHR principles							
	2-level modeling	Separation of EHR and demographic information	Service model	Version control—versioning	Version control—digital signature	Access control—access control list	Access control—configurations	Audit trailing
Method storage limitation	— ^a	X ^b	—	—	—	—	—	—
Integrity and confidentiality	—	—	—	X	X	X	X	X
Accountability	—	—	—	—	—	—	—	X
Record of processing	—	—	—	—	—	—	—	X
Availability of records of processing	—	—	—	—	—	—	—	X
Verification of the identity of the data subjects	—	X	—	—	—	—	—	—
Data subject access	—	—	—	—	—	X	X	—
Data subject direct access	—	—	X	—	—	X	X	—
Confirmation of data processing	—	—	—	X	—	—	—	X
Portability of personal data	X	—	—	—	—	—	—	—
Portability of personal data between controllers	X	—	—	—	—	—	—	—
Interoperability of systems and formats	X	—	X	—	—	—	—	—
Cross-border data transfers	X	—	—	—	—	—	—	—
Privacy by design	—	X	—	—	—	—	—	—
Privacy by default	—	X	—	—	X	X	X	—

^aRepresents no match.

^bX represents a match in the table.

General Data Protection Regulation Requirements Not Met by OpenEHR Principles

Textbox 2 presents the 35 GDPR requirements that are not met by openEHR principles.

Regarding requirements from group 3, although none of the openEHR principles identified could match this requirement, the implementation of openEHR architecture could help to fulfill the requirement related to explicit consent through the creation of an archetype *Consent*. This archetype would allow the

recording and management of Consent, allowing the controller to keep all the information necessary.

Regarding requirement 4.13, the information regarding the legitimate interest of processing could be included in an archetype, although even without this, the system can often infer legitimate access by analyzing, for example, hospital admission and discharge dates and association of subject to the general practitioner's clinic. Regardless, better methods are needed in the future. In case of consent being the legitimate interest for the processing, this information could be recorded along with the consent.

Textbox 2. Requirements that are not met by openEHR principles. DPIA: Data Protection Impact Assessment.

1. Limitations to data processing
 - Purpose limitation
 - Data minimization
 - Period of storage limitation
 - Limitation of processing of personal data
2. Data quality and accountability
 - Accuracy
 - Statement of accountability
 - Conditions of processing
 - Location of data
3. Consent by data subject
 - Explicit consent
 - Management of consent
 - Record of consent
 - Withdrawal of consent
 - Features of the consent
 - Lawfulness of processing after withdrawal of consent
 - Objection to processing
4. Data Subject empowerment
 - Information provided to data subject
 - Means to provide information to data subjects
 - Data subject request
 - Answer to request
 - Data subject request form
 - Information accessed by the data subject
 - Response to data subject request
 - Personal data rectification
 - Personal data erasure
 - Legitimate interest
5. Data breaches
 - Records of data breaches
 - Records of data breaches nature
 - Data breach description
 - Data breach notification deadline
 - Data breaches notification procedures
6. Data portability and interoperability
 - Communication between institutions
 - Cross border data transfers guarantee
7. Privacy control and impact assessment
 - Access control measures

- DPIA records preservation
- DPIA consultation

Discussion

Principal Findings

OpenEHR acts mainly on requirements that either shape the functional layer of the system or relate to data traceability, integrity, and confidentiality. Data protection by design, portability, and interoperability are ensured by openEHR's architecture because of the 2-level modeling and separation of clinical and demographic data. Personal data integrity and confidentiality are mainly addressed by the access control, versioning, and audit trail features.

Nevertheless, openEHR is a valuable tool for the fulfillment of requirements that are not directly met, such as definition of notification forms (for data subjects and data authorities), the creation of a means of communication for records, and preservation of Data Protection Impact Assessment and records of compliance with codes of conduct and certifications.

These requirements need to be addressed from an organizational point of view, either through the reform of existing processes or the definition of new ones. However, the versioning and audit trail features can support the recording of important information related to the data processing and data breaches.

Apart from the clinical content models (archetypes and templates), openEHR does not support the automatic creation of other needed documentation for GDPR such as a structure to store the consent but can be backed up by the principles (namely the traceability of the data, actions, and users).

OpenEHR's architectural features can still help fulfill requirements related to consent. Versioning and audit trailing allow systems to record and verify any action or access made in the EHR. By acknowledging the deadline for the processing,

it is possible to identify if there are any data being processed without the consent of the data subjects. Thus, even if the openEHR architecture does not include dedicated features for some GDPR requirements, it presents itself as an important support in relation to the processes that an institution must implement.

It should be noted that some of the GDPR requirements, namely the ones related to the organizational processes, are probably not satisfiable by any EHR architecture. However, it is important to note the organizational reforms that must be conducted require actions not only at the level of their organizational processes and services but also specifically at the level of their systems.

Limitations

To our knowledge, there is no formal list of GDPR requirements for an EHR system. The list of requirements we propose in this study is intended as a starting point for further discussion and future work.

Conclusions

OpenEHR is a promising approach to the development of EHR systems compliant with GDPR, allowing institutions to respond to functional needs focused on the privacy and security of health data. It is also a strong solution for issues related to data portability and data protection by default, which are now required by the regulation.

Primarily, openEHR is a good solution for issues related to privacy and data protection, the main goals of GDPR.

The use of IT has become essential to health care delivery. OpenEHR defines an integrated environment, focused on the provision of health care and access to quality information, which helps institutions conform to the GDPR requirements ensuring the privacy and protection of personal data.

Acknowledgments

The authors would like to acknowledge the project NanoSTIMA (NORTE-01-0145-FEDER-000016), which is financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund.

Conflicts of Interest

TB is one of the main authors of the openEHR specifications and is on occasion paid by the openEHR Foundation for R&D work relating to them.

References

1. Detmer DE, Steen EB, Dick RS, Institute of Medicine. The Computer-based Patient Record: An Essential Technology For Health Care, Revised Edition. Washington, DC: National Academies Press; 1997:209.
2. Reis ZS, Maia TA, Marcolino MS, Becerra-Posada F, Novillo-Ortiz D, Ribeiro AL. Is there evidence of cost benefits of electronic medical records, standards, or interoperability in hospital information systems? Overview of systematic reviews. *JMIR Med Inform* 2017 Aug 29;5(3):e26 [FREE Full text] [doi: [10.2196/medinform.7400](https://doi.org/10.2196/medinform.7400)] [Medline: [28851681](https://pubmed.ncbi.nlm.nih.gov/28851681/)]
3. Moro Visconti R. Big data-driven healthcare project financing. *SSRN Electron J* 2017 Mar 1. [doi: [10.2139/ssrn.2925790](https://doi.org/10.2139/ssrn.2925790)]

4. The Office of the National Coordinator for Health Information Technology. 2015 Apr. Guide to Privacy and Security of Health Information URL: <https://www.healthit.gov/sites/default/files/resources/privacy-and-security-guide.pdf> [accessed 2019-02-02] [WebCite Cache ID 75sWi6BNE]
5. Kulynych J, Greely HT. Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. *J Law Biosci* 2017 Jan 15;4(1):94-132 [FREE Full text] [doi: [10.1093/jlb/lsw061](https://doi.org/10.1093/jlb/lsw061)] [Medline: [28852559](https://pubmed.ncbi.nlm.nih.gov/28852559/)]
6. Great Britain National Audit Office. 2017. Investigation: WannaCry cyber attack and the NHS Internet URL: <https://www.nao.org.uk/report/investigation-wannacry-cyber-attack-and-the-nhs/> [accessed 2019-02-02] [WebCite Cache ID 75sWwvgyY]
7. HIMMS. 2018. 2018 HIMSS Cybersecurity Survey URL: <https://www.himss.org/2018-himss-cybersecurity-survey> [accessed 2019-02-02] [WebCite Cache ID 75sX2PWOP]
8. Williams P, Woodward A. Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Med Devices (Auckl)* 2015;8:305-316 [FREE Full text] [doi: [10.2147/MDER.S50048](https://doi.org/10.2147/MDER.S50048)] [Medline: [26229513](https://pubmed.ncbi.nlm.nih.gov/26229513/)]
9. Mense A, Blobel B. HL7 standards and components to support implementation of the European General Data Protection Regulation (GDPR). *Eur J Biomed Inform* 2017 Oct;13(1):33 [FREE Full text]
10. European Society of Radiology (ESR). The new EU General Data Protection Regulation: what the radiologist should know. *Insights Imaging* 2017 Jun;8(3):295-299 [FREE Full text] [doi: [10.1007/s13244-017-0552-7](https://doi.org/10.1007/s13244-017-0552-7)] [Medline: [28439718](https://pubmed.ncbi.nlm.nih.gov/28439718/)]
11. The European Council and The European Parliament. 2016 May 4. EUR-Lex. General Data Protection Regulation URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> [accessed 2019-02-02] [WebCite Cache ID 75sY6SKxi]
12. Rumbold JM, Pierscionek B. The effect of the general data protection regulation on medical research. *J Med Internet Res* 2017 Feb 24;19(2):e47 [FREE Full text] [doi: [10.2196/jmir.7108](https://doi.org/10.2196/jmir.7108)] [Medline: [28235748](https://pubmed.ncbi.nlm.nih.gov/28235748/)]
13. Beale T, Heard S. OpenEHR Foundation. 2007 Jan. openEHR architecture- architecture overview URL: https://specifications.openehr.org/releases/BASE/Release-1.0.3/architecture_overview.html [accessed 2019-02-02] [WebCite Cache ID 75sYkM6g7]
14. 1233-1998 - IEEE Guide for Developing System Requirements Specifications. Hoboken, USA: IEEE; Dec 29, 1998.
15. Molnár-Gábor F. Germany: a fair balance between scientific freedom and data subjects' rights? *Hum Genet* 2018 Aug;137(8):619-626 [FREE Full text] [doi: [10.1007/s00439-018-1912-1](https://doi.org/10.1007/s00439-018-1912-1)] [Medline: [30116955](https://pubmed.ncbi.nlm.nih.gov/30116955/)]
16. Phillips M. International data-sharing norms: from the OECD to the General Data Protection Regulation (GDPR). *Hum Genet* 2018 Aug;137(8):575-582 [FREE Full text] [doi: [10.1007/s00439-018-1919-7](https://doi.org/10.1007/s00439-018-1919-7)] [Medline: [30069638](https://pubmed.ncbi.nlm.nih.gov/30069638/)]
17. Chassang G. The impact of the EU general data protection regulation on scientific research. *Ecancermedicallscience* 2017 Jan 3;11:709 [FREE Full text] [doi: [10.3332/ecancer.2017.709](https://doi.org/10.3332/ecancer.2017.709)] [Medline: [28144283](https://pubmed.ncbi.nlm.nih.gov/28144283/)]
18. Gonçalves-Ferreira D, Leite M, Santos-Pereira C, Correia ME, Antunes L, Cruz-Correia R. HS.Register - an audit-trail tool to respond to the General Data Protection Regulation (GDPR). *Stud Health Technol Inform* 2018;247:81-85. [Medline: [29677927](https://pubmed.ncbi.nlm.nih.gov/29677927/)]
19. The openEHR Foundation. Implementation Technology Specifications (ITS) - Latest URL: <https://specifications.openehr.org/releases/ITS/latest/docs/index> [accessed 2019-02-02] [WebCite Cache ID 75sb1wwBB]
20. The openEHR Foundation. openEHR Platform Abstract Service Model URL: https://specifications.openehr.org/releases/SM/latest/openehr_platform.html [accessed 2019-02-02] [WebCite Cache ID 75sb6bAhu]
21. Pussewalage HS, Oleshchuk VA. Attribute based access control scheme with controlled access delegation for collaborative E-health environments. 2017 Dec Presented at: IEEE 18th Int Conf e-Health Networking, Appl Serv Internet IEEE; September 14-16, 2016; Munich, Germany p. 50-64. [doi: [10.1109/HealthCom.2016.7749516](https://doi.org/10.1109/HealthCom.2016.7749516)]
22. Moura P, Fazendeiro P, Marques P, Ferreira A. SoTRAAACE: Socio-technical risk-adaptable access control model. : IEEE; 2017 Presented at: 2017 International Carnahan Conference on Security Technology (ICCST); October 23-26, 2017; Madrid, Spain. [doi: [10.1109/CCST.2017.8167835](https://doi.org/10.1109/CCST.2017.8167835)]
23. The openEHR Foundation. Service Model (SM) - Latest URL: <https://specifications.openehr.org/releases/SM/latest/docs/index> [accessed 2019-02-02] [WebCite Cache ID 75sbZ6e8j]

Abbreviations

- EHR:** electronic health record
- DPIA:** Data Protection Impact Assessment
- EU:** European Union
- GDPR:** General Data Protection Regulation
- HIS:** hospital information system
- IS:** information system
- IT:** information technology

Edited by E Perakslis, M Stanley; submitted 15.01.18; peer-reviewed by D Moner Cano, I Mircheva; comments to author 29.03.18; revised version received 02.11.18; accepted 14.12.18; published 25.03.19.

Please cite as:

Gonçalves-Ferreira D, Sousa M, Bacelar-Silva GM, Frade S, Antunes LF, Beale T, Cruz-Correia R

OpenEHR and General Data Protection Regulation: Evaluation of Principles and Requirements

JMIR Med Inform 2019;7(1):e9845

URL: <http://medinform.jmir.org/2019/1/e9845/>

doi: [10.2196/medinform.9845](https://doi.org/10.2196/medinform.9845)

PMID: [30907730](https://pubmed.ncbi.nlm.nih.gov/30907730/)

©Duarte Gonçalves-Ferreira, Mariana Sousa, Gustavo M Bacelar-Silva, Samuel Frade, Luís Filipe Antunes, Thomas Beale, Ricardo Cruz-Correia. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Transcription Errors of Blood Glucose Values and Insulin Errors in an Intensive Care Unit: Secondary Data Analysis Toward Electronic Medical Record-Glucometer Interoperability

Azizeh Khaled Sowan¹, RN, MSDA, PhD; Ana Vera², RN, MSN; Ashwin Malshe³, PhD; Charles Reed², RN, CNRN, PhD

¹School of Nursing, University of Texas Health at San Antonio, San Antonio, TX, United States

²Center for Clinical Excellence, University Health System, San Antonio, TX, United States

³College of Business, University of Texas at San Antonio, San Antonio, TX, United States

Corresponding Author:

Azizeh Khaled Sowan, RN, MSDA, PhD
School of Nursing
University of Texas Health at San Antonio
7703 Floyd Curl Drive - MC 7975
San Antonio, TX, 78229
United States
Phone: 1 210 567 5799
Fax: 1 210 567 1719
Email: sowan@uthscsa.edu

Abstract

Background: Critically ill patients require constant point-of-care blood glucose testing to guide insulin-related decisions. Transcribing these values from glucometers into a paper log and the electronic medical record is very common yet error-prone in intensive care units, given the lack of connectivity between glucometers and the electronic medical record in many US hospitals.

Objective: We examined (1) transcription errors of glucometer blood glucose values documented in the paper log and in the electronic medical record vital signs flow sheet in a surgical trauma intensive care unit, (2) insulin errors resulting from transcription errors, (3) lack of documenting these values in the paper log and the electronic medical record vital signs flow sheet, and (4) average time for docking the glucometer.

Methods: This secondary data analysis examined 5049 point-of-care blood glucose tests. We obtained values of blood glucose tests from bidirectional interface software that transfers the meters' data to the electronic medical record, the paper log, and the vital signs flow sheet. We obtained patient demographic and clinical-related information from the electronic medical record.

Results: Of the 5049 blood glucose tests, which were pertinent to 234 patients, the total numbers of undocumented or untranscribed tests were 608 (12.04%) in the paper log, 2064 (40.88%) in the flow sheet, and 239 (4.73%) in both. The numbers of transcription errors for the documented tests were 98 (2.21% of 4441 documented tests) in the paper log, 242 (8.11% of 2985 tests) in the flow sheet, and 43 (1.64% of 2616 tests) in both. The numbers of transcription errors per patient were 0.4 (98 errors/234 patients) in the paper log, 1 (242 errors/234 patients) in the flow sheet, and 0.2 in both (43 errors/234 patients). Transcription errors in the paper log, the flow sheet, and in both resulted in 8, 24, and 2 insulin errors, respectively. As a consequence, patients were given a lower or higher insulin dose than the dose they should have received had there been no errors. Discrepancies in insulin doses were 2 to 8 U lower doses in paper log transcription errors, 10 U lower to 3 U higher doses in flow sheet transcription errors, and 2 U lower in transcription errors in both. Overall, 30 unique insulin errors affected 25 of 234 patients (10.7%). The average time from point-of-care testing to meter docking was 8 hours (median 5.5 hours), with some taking 56 hours (2.3 days) to be uploaded.

Conclusions: Given the high dependence on glucometers for point-of-care blood glucose testing in intensive care units, full electronic medical record-glucometer interoperability is required for complete, accurate, and timely documentation of blood glucose values and elimination of transcription errors and the subsequent insulin-related errors in intensive care units.

(*JMIR Med Inform* 2019;7(1):e11873) doi:[10.2196/11873](https://doi.org/10.2196/11873)

KEYWORDS

transcription errors; blood glucose; insulin errors; interoperability; glucometer; electronic medical records; secondary data analysis; intensive care units; medication errors

Introduction

Background

Glycemic control in critically ill patients is essential to improve clinical outcomes and decrease morbidity and mortality [1-8], specifically for patients admitted to intensive care units (ICUs) for more than 3 days [2] and for patients admitted to surgical trauma ICUs (STICUs) compared with medical ICUs [7]. Critically ill patients require constant point-of-care tests (POCTs) for blood glucose to guide initiation and titration decisions regarding continuous insulin infusion following insulin management protocols. Handheld blood glucose monitoring devices or glucometers are widely used in ICUs for this purpose for convenience and portability [9,10].

Transcribing blood glucose readings from glucometers into a paper log and different flow sheets in the electronic medical record (EMR) by health care professionals is a very common yet error-prone practice in ICUs, given the lack of interoperability or connectivity between glucometers and the EMR in many US hospitals [11]. Interoperability allows for wireless transfer of blood glucose values from glucometers to the EMR without the need for manual data entry. Despite the call for system interoperability and emerging research describing frameworks and prototypes for seamless integration of medical device data into the EMR using different connectivity standards [12-15], medical device-EMR connectivity is limited in the United States. In a national survey of 825 US hospitals, the Health Information and Management Systems Society Analytics team reported a lack of any interface between EMRs and medical devices in 70% of the hospitals. The remaining 30% of the hospitals reported an interface of an average of 2.6 device types (out of 11 devices) with their EMRs. Interestingly, none of the hospitals provided an interface between glucometers and the EMR [11].

Extensive literature exists on the use of glucometers in ICUs. However, most studies focused on the glucometers' accuracy in comparison with other blood glucose analytical measures [16-24]. Research on transcription errors is also available [25-27]; however, there is a paucity of research on transcription errors of blood glucose values obtained by glucometers into the EMRs and the subsequent insulin errors [28]. Although the use of glucometers with high specificity and sensitivity is essential in critical care settings to prevent harmful effects of erroneous blood glucose readings and the subsequent underdose or overdose of insulin therapy, accurate and instant documentation of blood glucose values obtained by glucometers into the EMR is equally important to inform glycemic control and insulin management decisions.

Objective

This study examined (1) transcription errors of blood glucose values obtained by a glucometer that were documented in the paper log by technicians and in the EMR vital signs flow sheet

by nurses in the ICU, (2) insulin errors resulting from transcription errors of blood glucose values, (3) lack of documenting blood glucose values in the paper log and the EMR vital signs flow sheet, and (4) meter docking time.

Methods

Design, Sample, and Setting

This secondary data analysis study examined 5049 blood glucose tests for transcription errors, insulin errors, lack of documenting blood glucose values in the paper log and the EMR, and meter docking time. The study took place in a 30-bed STICU located in a 705-bed university teaching hospital with a large referral base in the southwestern United States. The STICU has an annual admission rate of 1600 patients and an approximate monthly admission rate of 133 patients. At the time of the study, there were 46 full-time and 11 part-time nurses and 13 technicians working in the unit. The average range of blood glucose POCTs performed on patients in the unit is 4200 to 4300 tests per month.

After obtaining institutional review board approval from the University of Texas and the University Health System (number 20140330H), we performed the audit of blood glucose tests and insulin data in a 20% stratified sample of all blood glucose tests available in the meters for patients admitted during 4 months (July to October 2016). Stratification was based on the working shift (day or night) as the only possible factor that may introduce transcription errors of blood glucose readings as a result of fatigue expected at the end of each working shift and on the night shift. Additionally, when we selected a blood glucose test, we also included all blood glucose tests pertinent to the same patient within the same episode of admission to evaluate errors per patient. This resulted in a total of 5049 blood glucose tests.

Description of the Point-of-Care Testing of Blood Glucose

The point-of-care glucose testing device is Accu-Chek Inform II (Roche Diagnostics Corporation, Indianapolis, IN, USA). Figure 1 depicts a functional workflow model for this process of testing. The process starts by the physician ordering a POCT. The nurse informs the technician about the order, who in turn performs the test using the glucometer and transcribes the result into a paper log—a grid that includes the patient's name, visit identification number (VIN), room number, time and date of the test, and the result. The VIN is a unique number for each patient episode of admission that is obtained by scanning the patient's wristband at the time of performing the test.

Nurses then manually enter the readings for each patient into the EMR vital signs flow sheet and use this information to inform their insulin management decisions following physician orders and insulin management protocols. Clinical decisions include whether to continue to monitor, repeat the test to verify critical blood glucose values, inform the physician, give insulin, and titrate the insulin drip based on the insulin management

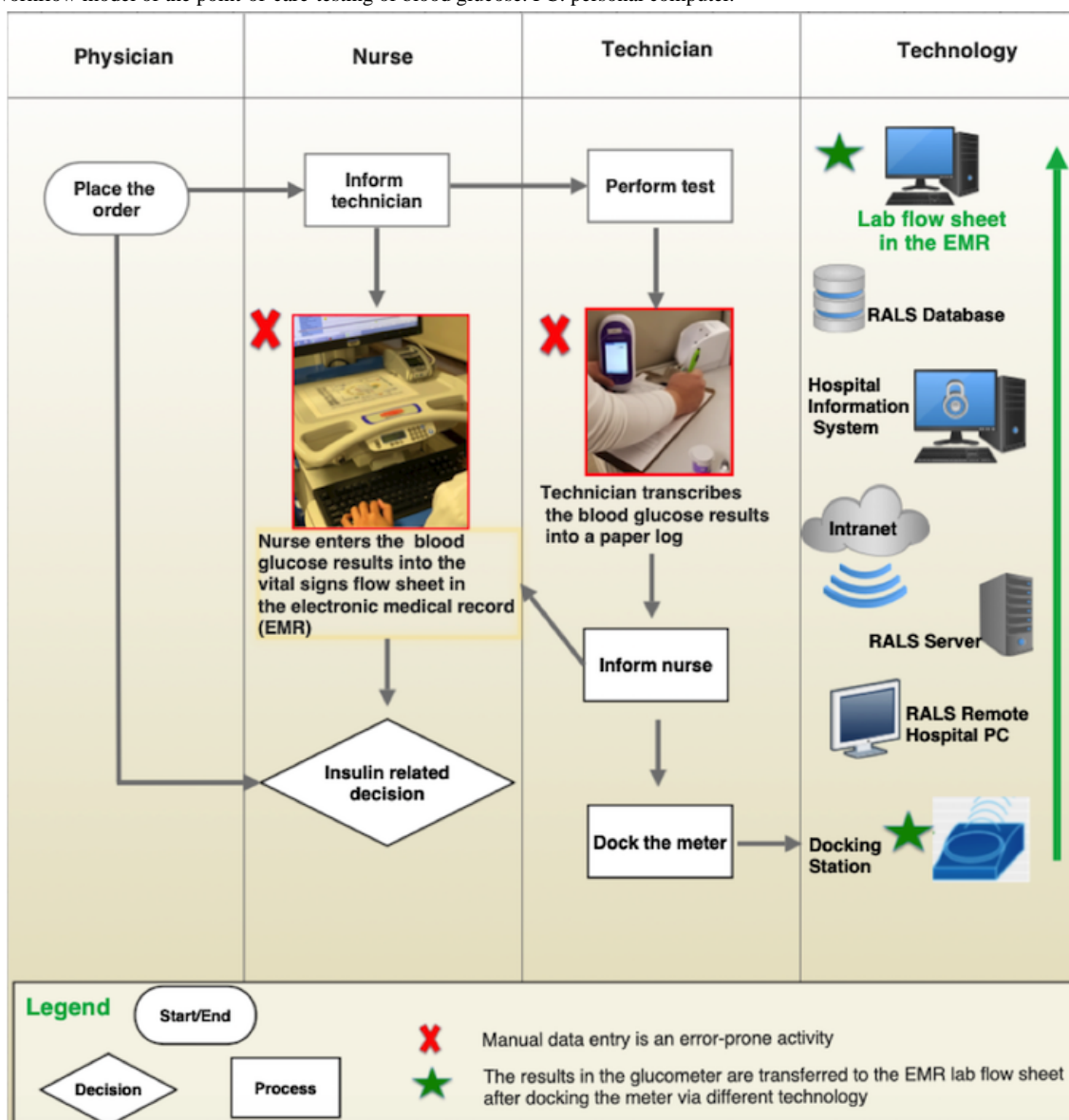
protocol. The blood glucose values entered by nurses into the EMR vital signs flow sheet can be obtained (1) from the glucometer itself by manually searching the readings using the time of the test and the patient VIN to locate the test value, (2) from the technician, who verbally endorses the value to the nurse if he or she is available in the unit, or (3) by checking the value transcribed by the technician into the paper log.

The technician docks the meter by placing it into the meter base unit within 24 hours after the time of the first test for a given day. Meters maintain log data for up to 2000 readings. Since the meter can be docked after 24 hours of use, nurses usually base their insulin management decisions on the readings transcribed by the technicians into the paper log or the readings entered by the nurses into the vital signs flow sheet. By docking the meter, readings are automatically uploaded into the RALS-Plus database, which interfaces with the EMR laboratory

flow sheet. These data include the examiner's employee identification number, patient identification (name, VIN), date and time of the test, time the meter was docked, and blood glucose values. It is worth noting that there is no direct link or seamless transfer of data in the EMR between the vital signs flow sheet and the EMR laboratory flow sheet.

The RALS-Plus v1.5.1 (Alere North America, LLC, Orlando, FL, USA) is a bidirectional interface software for in-hospital glucometers that uploads meter data into the EMR laboratory flow sheet only. The software also generates different types of reports for quality improvement. Data can be generated based on criteria such as the start and end date of the test, blood glucose values, patient VIN, sample type, and test location. Reports can be emailed, printed, saved, or exported into an Excel, rich text (rtf), or pdf file format.

Figure 1. Workflow model of the point-of-care-testing of blood glucose. PC: personal computer.



Main Outcome Variables

Transcription Errors

Since the focus of this study was transcription errors, we assumed that technicians follow best practices in obtaining blood samples and in meter use according to the unit policies and procedures and the glucometer's user manual. Blood glucose values uploaded into RALS-Plus are those in the meters and they are transferred to the EMR laboratory flow sheet. These

values are accurate. Transcribing blood glucose values from the meters to the paper log and the EMR vital signs flow sheet may result in 3 potential types of errors (Table 1). The "corresponding values" (Table 1) in the paper log and the EMR vital signs flow sheet are based on the same patient VIN, same date of the test, and within a 1-hour time frame from the POCT (time in RALS) to the time the test was transcribed into the paper log or the EMR vital signs flow sheet.

Table 1. Types of errors in transcribing blood glucose values from meters to the paper log and the electronic medical record (EMR) vital signs flow sheet.

Paper log	EMR vital signs flow sheet	
	Flow sheet correct	Flow sheet wrong
Paper correct	<i>No error:</i> The blood glucose value in the RALS database matches the corresponding value transcribed by technicians and nurses into the paper log and the EMR vital signs flow sheet for a given test in a given date and time.	<i>Vital signs flow sheet error:</i> Any discrepancy regardless of the magnitude between blood glucose value in the RALS database and the corresponding value transcribed by nurses into the EMR vital signs flow sheet.
Paper wrong	<i>Paper log error:</i> Any discrepancy regardless of the magnitude between blood glucose value in the RALS database and the corresponding value transcribed by the technician into the paper log.	<i>Paper log and vital signs flow sheet error:</i> The 2 blood glucose values transcribed by technicians and nurses into the paper log and the EMR vital signs flow sheet for a given test in a given date and time do not match the value in the RALS database.

Undocumented Values of Blood Glucose Tests

Untranscribed or undocumented blood glucose values are those available in RALS database but were not transcribed into the paper log or entered into the EMR vital signs flow sheet.

Insulin Errors Related to Erroneously Transcribed Blood Glucose Values

For each transcription error, we also examined whether that error resulted in giving the wrong dose of insulin. We evaluated the wrong insulin dose based on administering a higher or lower insulin dose, regardless of the magnitude of the difference, than the one recommended by the protocol for the correct blood glucose value (the value in the RALS system) or not giving insulin when it should be administered to the patient according to the insulin management protocol based on the correct blood glucose value.

Meter Docking Time

As mentioned above, we considered a 1-hour time frame from the POCT (time in RALS or glucometer) to the time the test result was transcribed into the paper log or the EMR vital signs flow sheet when we retrieved the time for transcribing blood glucose values. Meter docking time was retrieved from the RALS database and is the time from the POCT to the time meters were docked (readings were uploaded into the EMR laboratory flow sheet).

In addition to these outcomes, we also collected patient demographics and clinical-related information such as age, sex, diagnosis, diabetes status, admission and discharge dates, and total number of POCTs the patient underwent during the ICU stay.

Data Collection Procedure

We took the following steps in the sequence identified to collect the data. Three nurse educators collected the data from the paper log and the EMR vital signs flow sheet to enhance objectivity.

First, we accessed the RALS database for the selected study months and downloaded the Excel file (Microsoft Corporation). The file included the patient's name, VIN, EMR number, test date and time, blood glucose value (meter value), and time of docking the meter.

Second, we selected a stratified sample of 20% of the blood glucose readings and the related information from RALS from the Excel file. In addition to the 20% sample of readings, we went back and selected all pertinent blood glucose tests within the episode of admission for all VINs included in the stratified sample.

Third, for each test selected from RALS, we accessed the EMR and obtained patient demographics and clinical-related information based on the VIN, as well as the corresponding values of blood glucose transcribed into the vital signs flow sheet and time of documentation. We also accessed the laboratory flow sheet to make sure that the tests in RALS were pertinent to that patient.

Fourth, for each test selected from RALS (step 2) for each patient and based on the VIN, we accessed the paper log using the patient's name and VIN as the identifiers. We obtained the corresponding blood glucose value for each test using the date and a 1-hour time frame from the POCT (time in RALS) as the matching codes. We also obtained the actual time of the test documented in the paper log.

Data Analysis

We used R statistical computing software v3.5.1 (R Foundation) to analyze the data. Patients' characteristics and all types of errors were presented using descriptive statistics. We examined

the difference in number of POCTs between diabetic and nondiabetic patients using Student *t* test with significance set at $P < .05$.

We limited the analysis of transcription errors to cases where the results of the blood glucose tests were transcribed by clinicians and nurses. For example, the denominator for the paper log transcription errors was the number of blood glucose readings transcribed into the paper log, excluding missing values (ie, when the readings were not transcribed).

Results

Patient Characteristics and Number of Point-of-Care Tests

The 5049 blood glucose tests analyzed for transcription errors, undocumented blood glucose readings, and meter docking time were pertinent to 234 unique patients, each with a unique VIN. Table 2 presents the patients' characteristics. Most of the patients with documented diabetes status in the dataset did not have diabetes (93/234). Of the 234 patients, 97 were with unknown diabetes status. The average number of POCTs performed on diabetic patients (Table 3) was significantly higher than on nondiabetic patients ($t_{47} = -2.17$, $P = .03$). One of the patients had 792 POCTs during his stay (Table 3). The median number of POCTs for diabetic patients was 12 tests.

Table 2. Patient characteristics (N=234).

Patient characteristic	Statistics
Age in years, mean (SD)	57.5 (17.4)
Length of stay in days, mean (SD)	24.8 (48.3)
Number of point-of-care tests per patient, mean (SD)	25.5 (67.9)
Male sex, n (%)	131 (56.0)
Diabetes status, n (%)	
Yes	44 (32.1)
No	93 (67.9)
Missing	97 (41.5)

Table 3. Comparison of the number of point-of-care tests between diabetic and nondiabetic patients.

Diabetes status	Minimum	Median	Mean (SD)	Maximum
Diabetes (n=44)	1	12	60 (126)	792
No diabetes (n=93)	1	6	19 (40)	344

Missing Documentation and Transcription Errors

Table 4 describes the number of tests that were not transcribed into the paper log or the EMR vital signs flow sheet, or both, as well as the number of transcription errors. In the vital signs

flow sheet, 40.88% of the tests (2064/5049 tests) were not transcribed. Of the blood glucose tests, 4.73% (239/5049 tests) were not transcribed in the paper log and in the EMR vital signs flow sheet at the same time.

Table 4. Number of undocumented blood glucose tests and number of transcription errors among the 5049 tests analyzed.

Source	Undocumented tests, n (%)	Tests analyzed for errors, n	Errors among tests analyzed, n (%)	Range of error, mg/dL (mmol/L) ^a	Error rate per patient (N=234)
Paper log	608 (12.04)	4441	98 (2.21)	-92 to 92 (-5.1 to 5.1)	0.4 (98/234)
Vital signs flow sheet	2064 (40.88)	2985	242 (8.11)	-110 to 80 (-6.1 to 4.4)	1.0 (242/234)
Both	239 (4.73)	2616	43 (1.64)	N/A ^b	0.2 (43/234)

^aRange of difference between the correct blood glucose value and the erroneously transcribed value.

^bN/A: not applicable.

We analyzed all types of transcription errors when the blood glucose value was transcribed (Table 4). Of the transcription errors among the 4441 transcribed tests in the paper logs, there were 98 (2.21%) errors. These errors were related to 30 of the

234 patients (12.8%). Of the 2985 transcribed values in the vital signs flow sheet, there were 242 (8.11%) errors related to 63 of the 234 patients (26.9%). The total number of paper log and vital signs flow sheet transcription errors among the 2616 tests

analyzed was 43 (1.64%), related to 24 of the 234 patients (10.3%). Overall, among the 234 patients, there were 68 (29.1%) unique patients involved in all types of errors.

Errors in the paper log resulted in transcribing a blood glucose value that was up to 92 mg/dL (5.1 mmol/L) lower or 92 mg/dL (5.1 mmol/L) higher than the correct value (the one in the EMR laboratory flow sheet or RALS). However, most errors, those between the 25th and 75th percentiles, were 12 mg/dL (0.7 mmol/L) lower to 7 mg/dL (0.4 mmol/L) higher than the

accurate value. In the EMR vital signs flow sheet, the difference between the correct blood glucose value and the erroneously transcribed value was 110 mg/dL (6.1 mmol/L) lower to 80 mg/dL (4.4 mmol/L) higher. Most errors, those between the 25th and 75th percentiles, were 3 mg/dL (0.16 mmol/L) lower to 4 mg/dL (0.2 mmol/L) higher than the accurate values.

There were no significant differences in the number of transcription errors between the day shift and night shift (Table 5).

Table 5. Difference in transcription errors between the day shift and night shift.

Source	Total errors, n (%)	Day shift, n (%)	Night shift, n (%)	Chi-square	df	P value
Paper log	98/4441 (2.21)	53/2790 (1.90)	45/1651 (2.73)	2.9	1	.09
Vital signs flow sheet	242/2985 (8.11)	163/1847 (8.83)	79/1138 (6.94)	3.1	1	.08
Both	43/2616 (1.64)	24/1639 (1.46)	19/977 (1.94)	0.6	1	.44

Insulin Errors

The 242 transcription errors in the EMR vital signs flow sheet resulted in 24 insulin errors. These errors resulted in giving 10 U lower to 3 U higher insulin dose than the dose that should have been given had there been no transcription errors. The 98 transcription errors in the paper log resulted in 8 insulin errors and giving 2 to 8 U lower insulin dose than the dose that should have been given had there been no transcription errors. The 43 errors in the EMR vital signs flow sheet and paper logs resulted in 2 insulin errors, both with 2 U lower than the correct insulin dose. Overall, there were 30 unique insulin errors that affected 25 of the 234 patients (10.7%).

Documentation Time

The average time from the POCT to the time meters were docked (readings were uploaded into the EMR laboratory flow sheet) was 8 hours with a median of 5.5 hours. Most readings, between the first and the third quartiles, took 1.3 to 12 hours to be uploaded into the EMR laboratory flow sheet. Some of the readings took 56 hours (2.3 days) to be uploaded into the EMR laboratory flow sheet.

In addition to these outcomes, we found 40 readings that were documented to some patients' EMRs and the paper log after the date of discharge.

Discussion

Principal Findings

This study examined transcription errors of blood glucose tests obtained by a glucometer and documented in the paper log by ICU technicians and in the EMR vital signs flow sheet by ICU nurses. Insulin errors resulted from transcription errors of blood glucose values, the number of undocumented blood glucose tests in the paper log and the EMR vital signs flow sheet, and the average meter docking time. Research on the use of glucometers in ICU and non-ICU settings is extensive. However, most of these studies focused on precision and accuracy of the glucometers, sources of glucometer measurement errors, and the difference in sensitivity and specificity between glucometer devices from different vendors [16-29]. Nevertheless,

glucometers are commonly used handheld devices to measure blood glucose at the point of care, specifically in ICUs to inform timely clinical decisions regarding insulin therapy. To our knowledge, this is the first study to examine transcription errors of blood glucose tests obtained by glucometers and to focus on the urgent need for EMR-glucometer interoperability.

Transcription errors ranged from 2% for paper log errors to 8% for vital signs flow sheet errors. These errors resulted in a total of 30 insulin errors and affected 11% of the patients. The higher percentage of transcription errors in the vital signs flow sheet than in the paper log might be explained by a clinical workflow that has nurses obtain the results of blood glucose tests from 3 different sources, which are the paper log, the technicians, or the glucometers, while the technicians obtain the values only from the glucometers. Transcription errors in the vital signs flow sheet are clinically more significant than transcription errors in the paper log because they inform nurses' insulin management decisions. These errors affected 63 (27%) patients.

It is important to note that we examined transcription errors and the associated insulin errors only when the blood glucose test results were transcribed by technicians and nurses. The very high percentage of untranscribed values (ie, up to 41% untranscribed into the vital signs flow sheet, n=2064) could mask the actual rate of transcription errors. Possible explanations for not transcribing blood glucose values might be workload issues and the assumption that all readings eventually will be available in the laboratory flow sheet in the EMR after docking the meter. In addition, finding 40 readings documented to some patients' EMRs and the paper log after the date of discharge is alarming. This means that technicians were not scanning the patient bracelet but probably a sticker that remained on the paper log or the patient monitor or bed. Although eliminating the use of a paper log via a full EMR-glucometer interoperability could decrease this error, adherence to the unit policies and procedures for safe testing is critical for complete elimination of this error.

Although a partial interface exists in our hospital between glucometers and the EMR through the RALS bidirectional interface software, this interface transfers the data only to the EMR central laboratory flow sheet. Additionally, based on the unit policies and procedures, the meters should be docked within

24 hours by technicians. This long time period hinders the availability of the tests' values at the point of care, making these data unusable for immediate clinical decisions. Furthermore, our results showed that, in reality, docking the meters might take up to more than 2 days. Therefore, there is an urgent need for full glucometer-EMR connectivity to allow for seamless transfer of meter data into other fields of the EMR (ie, the vital signs flow sheet) in order to eliminate data transcription errors and the associated insulin errors.

The few available studies on medical devices-EMR connectivity have focused on vital signs monitors in ICUs and supported improved efficiency and elimination of transcription errors when vital signs monitoring devices are connected to the EMR [14]. The results of our study support the urgent need for a comprehensive and instant connectivity to transfer glucometer data to all fields of the EMR to better inform clinical decisions and eliminate insulin errors associated with transcription errors. On the other hand, from an engineering perspective, interoperability challenges do exist. These may include lack of research describing successes and challenges, the complexity of data elements, and the difference in type of information and formats in which information is stored and displayed. Most important, studies supported the potential for new types of errors in device connectivity, such as transferring the data into the wrong patient's EMR, in addition to the slow speed of the interface attributed to the slow speed of older medical devices and computers [12]. Therefore, the process of and errors associated with interoperability should be carefully examined.

Limitations

The results of this study should be interpreted in light of the following limitations. First, since workload, admission rate, and the large number of monthly POCTs are inherent factors that may affect transcription errors, our results can only be generalized to STICUs with a similar workload and rate of POCTs. Second, we limited the errors examined in this study to transcription errors; measurement errors of blood glucose values that may result from inappropriate testing or scanning the wrong patients were beyond the scope of this study. Third, because we collected retrospective data, our risk assessment was limited to identifying the number of insulin errors resulting from transcription errors without identifying the clinical consequences or adverse events of insulin errors. On the other hand, insulin is a high-alert medication and errors in its administration may cause serious hypoglycemia and hyperglycemia, seizures, coma, ketoacidosis, and even death [30].

Conclusions

Transcription errors of blood glucose values obtained by glucometers do exist and result in insulin errors. Given the high dependence on glucometers for POCTs of blood glucose in ICUs, full EMR-glucometer interoperability is required for complete and accurate documentation of blood glucose values, and elimination of transcription errors and the subsequent insulin-related errors in ICUs.

Acknowledgments

This project was funded by the University Health System's Center for Clinical Excellence and the University of Texas Health School of Nursing. The role of the funding sources was limited to providing financial support for the conduct of the research. Sponsors were not involved in study design; collection, analysis, or interpretation of data; writing of the manuscript; or the decision to submit the article for publication.

Conflicts of Interest

None declared.

References

1. Krinsley JS, Maurer P, Holewinski S, Hayes R, McComsey D, Umpierrez GE, et al. Glucose control, diabetes status, and mortality in critically ill patients: the continuum from intensive care unit admission to hospital discharge. *Mayo Clin Proc* 2017 Dec;92(7):1019-1029. [doi: [10.1016/j.mayocp.2017.04.015](https://doi.org/10.1016/j.mayocp.2017.04.015)] [Medline: [28645517](https://pubmed.ncbi.nlm.nih.gov/28645517/)]
2. Van den Berghe G, Wilmer A, Hermans G, Meersseman W, Wouters PJ, Milants I, et al. Intensive insulin therapy in the medical ICU. *N Engl J Med* 2006 Feb 02;354(5):449-461. [doi: [10.1056/NEJMoa052521](https://doi.org/10.1056/NEJMoa052521)] [Medline: [16452557](https://pubmed.ncbi.nlm.nih.gov/16452557/)]
3. Silva-Perez LJ, Benitez-Lopez MA, Varon J, Surani S. Management of critically ill patients with diabetes. *World J Diabetes* 2017 Mar 15;8(3):89-96 [FREE Full text] [doi: [10.4239/wjd.v8.i3.89](https://doi.org/10.4239/wjd.v8.i3.89)] [Medline: [28344751](https://pubmed.ncbi.nlm.nih.gov/28344751/)]
4. Williams JB, Peterson ED, Albrecht A, Li S, Hirji SA, Ferguson T, et al. Glycemic control in patients undergoing coronary artery bypass graft surgery: clinical features, predictors, and outcomes. *J Crit Care* 2017 Dec;42:328-333. [doi: [10.1016/j.jcrc.2017.09.013](https://doi.org/10.1016/j.jcrc.2017.09.013)] [Medline: [28935429](https://pubmed.ncbi.nlm.nih.gov/28935429/)]
5. Murphy CV, Coffey R, Cook CH, Gerlach AT, Miller SF. Early glycemic control in critically ill patients with burn injury. *J Burn Care Res* 2011;32(6):583-590. [doi: [10.1097/BCR.0b013e31822dc3da](https://doi.org/10.1097/BCR.0b013e31822dc3da)] [Medline: [21841493](https://pubmed.ncbi.nlm.nih.gov/21841493/)]
6. Krinsley JS. Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients. *Mayo Clin Proc* 2003 Dec;78(12):1471-1478. [doi: [10.4065/78.12.1471](https://doi.org/10.4065/78.12.1471)] [Medline: [14661676](https://pubmed.ncbi.nlm.nih.gov/14661676/)]
7. Griesdale DEG, de Souza RJ, van Dam RM, Heyland DK, Cook DJ, Malhotra A, et al. Intensive insulin therapy and mortality among critically ill patients: a meta-analysis including NICE-SUGAR study data. *CMAJ* 2009 Apr 14;180(8):821-827 [FREE Full text] [doi: [10.1503/cmaj.090206](https://doi.org/10.1503/cmaj.090206)] [Medline: [19318387](https://pubmed.ncbi.nlm.nih.gov/19318387/)]

8. Wiener RS, Wiener DC, Larson RJ. Benefits and risks of tight glucose control in critically ill adults: a meta-analysis. *JAMA* 2008 Aug 27;300(8):933-944. [doi: [10.1001/jama.300.8.933](https://doi.org/10.1001/jama.300.8.933)] [Medline: [18728267](https://pubmed.ncbi.nlm.nih.gov/18728267/)]
9. Tonyushkina K, Nichols JH. Glucose meters: a review of technical challenges to obtaining accurate results. *J Diabetes Sci Technol* 2009 Jul 01;3(4):971-980 [FREE Full text] [doi: [10.1177/193229680900300446](https://doi.org/10.1177/193229680900300446)] [Medline: [20144348](https://pubmed.ncbi.nlm.nih.gov/20144348/)]
10. Klonoff DC. Point-of-care blood glucose meter accuracy in the hospital setting. *Diabetes Spectr* 2014 Aug;27(3):174-179 [FREE Full text] [doi: [10.2337/diaspect.27.3.174](https://doi.org/10.2337/diaspect.27.3.174)] [Medline: [26246776](https://pubmed.ncbi.nlm.nih.gov/26246776/)]
11. HIMSS. Medical devices landscape: current and future adoption, integration with EMRs, and connectivity. A HIMSS Analytics white paper. Chicago, IL: HIMSS Analytics; 2010. URL: https://www.lantronix.com/wp-content/uploads/pdf/Medical-Devices-Landscape_Lantronix_HIMMS_WP.pdf [accessed 2018-08-06] [WebCite Cache ID 71UrMwpmC]
12. Choi JS, Lee JH, Park JH, Nam HS, Kwon H, Kim D, et al. Design and implementation of a seamless and comprehensive integrated medical device interface system for outpatient electronic medical records in a general hospital. *Int J Med Inform* 2011 Apr;80(4):274-285. [doi: [10.1016/j.ijmedinf.2010.11.007](https://doi.org/10.1016/j.ijmedinf.2010.11.007)] [Medline: [21134783](https://pubmed.ncbi.nlm.nih.gov/21134783/)]
13. Medical device plug-and-play (MD PnP) interoperability program. Cambridge, MA: MD PnP; 2012 Oct. URL: http://www.mdnp.org/uploads/Oct12_MD_PnP_White_Paper.pdf [accessed 2018-08-06] [WebCite Cache ID 71UrTKDRT]
14. Fonseca T, Ribeiro C, Granja C. Vital signs in intensive care: automatic acquisition and consolidation into electronic patient records. *J Med Syst* 2009 Feb;33(1):47-57. [Medline: [19238896](https://pubmed.ncbi.nlm.nih.gov/19238896/)]
15. Mense A, Sauermann S, Gerbovic G, Frohner M, Pohn B, Bruckner R, et al. Usage of international standards for integrating extramural monitoring and personal health device data into medical information infrastructure. *Stud Health Technol Inform* 2010;160(Pt 1):550-553. [Medline: [20841747](https://pubmed.ncbi.nlm.nih.gov/20841747/)]
16. Mahmoodpoor A, Hamishehkar H, Shadvar K, Sanaie S, Iranpour A, Fattahi V. Validity of bedside blood glucose measurement in critically ill patients with intensive insulin therapy. *Indian J Crit Care Med* 2016 Nov;20(11):653-657 [FREE Full text] [doi: [10.4103/0972-5229.194009](https://doi.org/10.4103/0972-5229.194009)] [Medline: [27994380](https://pubmed.ncbi.nlm.nih.gov/27994380/)]
17. Zhang R, Isakow W, Kollef MH, Scott MG. Performance of a modern glucose meter in ICU and general hospital inpatients: 3 years of real-world paired meter and central laboratory results. *Crit Care Med* 2017 Sep;45(9):1509-1514. [doi: [10.1097/CCM.0000000000002572](https://doi.org/10.1097/CCM.0000000000002572)] [Medline: [28640025](https://pubmed.ncbi.nlm.nih.gov/28640025/)]
18. DuBois JA, Slingerland RJ, Fokkert M, Roman A, Tran NK, Clarke W, et al. Bedside glucose monitoring-is it safe? A new, regulatory-compliant risk assessment evaluation protocol in critically ill patient care settings. *Crit Care Med* 2017 Apr;45(4):567-574 [FREE Full text] [doi: [10.1097/CCM.0000000000002252](https://doi.org/10.1097/CCM.0000000000002252)] [Medline: [28169943](https://pubmed.ncbi.nlm.nih.gov/28169943/)]
19. Joseph JI. Analysis: new point-of-care blood glucose monitoring system for the hospital demonstrates satisfactory analytical accuracy using blood from critically ill patients--an important step toward improved blood glucose control in the hospital. *J Diabetes Sci Technol* 2013 Sep 01;7(5):1288-1293 [FREE Full text] [doi: [10.1177/193229681300700519](https://doi.org/10.1177/193229681300700519)] [Medline: [24124956](https://pubmed.ncbi.nlm.nih.gov/24124956/)]
20. Inoue S, Egi M, Kotani J, Morita K. Accuracy of blood-glucose measurements using glucose meters and arterial blood gas analyzers in critically ill adult patients: systematic review. *Crit Care* 2013 Mar 18;17(2):R48 [FREE Full text] [doi: [10.1186/cc12567](https://doi.org/10.1186/cc12567)] [Medline: [23506841](https://pubmed.ncbi.nlm.nih.gov/23506841/)]
21. Slater-MacLean L, Cembrowski G, Chin D, Shalapay C, Binette T, Hegadoren K, et al. Accuracy of glycemic measurements in the critically ill. *Diabetes Technol Ther* 2008 Jun;10(3):169-177. [doi: [10.1089/dia.2008.0263](https://doi.org/10.1089/dia.2008.0263)] [Medline: [18473690](https://pubmed.ncbi.nlm.nih.gov/18473690/)]
22. Mraovic B, Schwenk ES, Epstein RH. Intraoperative accuracy of a point-of-care glucose meter compared with simultaneous central laboratory measurements. *J Diabetes Sci Technol* 2012 May 01;6(3):541-546 [FREE Full text] [doi: [10.1177/193229681200600308](https://doi.org/10.1177/193229681200600308)] [Medline: [22768884](https://pubmed.ncbi.nlm.nih.gov/22768884/)]
23. Hoedemaekers CWE, Klein Gunnewiek JMT, Prinsen MA, Willems JL, Van der Hoeven JG. Accuracy of bedside glucose measurement from three glucometers in critically ill patients. *Crit Care Med* 2008 Nov;36(11):3062-3066. [doi: [10.1097/CCM.0b013e318186ffe6](https://doi.org/10.1097/CCM.0b013e318186ffe6)] [Medline: [18824915](https://pubmed.ncbi.nlm.nih.gov/18824915/)]
24. American Association of Diabetic Educators. Practice advisory blood glucose meter accuracy. Chicago, IL: AADE; 2013 Sep 16. URL: https://www.diabeteseducator.org/docs/default-source/legacy-docs/_resources/pdf/research/Practice_Advisory_BGM_FINAL.pdf [accessed 2018-08-07] [WebCite Cache ID 71UrileDO]
25. MacKay M, Anderson C, Boehme S, Cash J, Zobell J. Frequency and severity of parenteral nutrition medication errors at a large children's hospital after implementation of electronic ordering and compounding. *Nutr Clin Pract* 2016 Apr;31(2):195-206. [doi: [10.1177/0884533615591606](https://doi.org/10.1177/0884533615591606)] [Medline: [26214511](https://pubmed.ncbi.nlm.nih.gov/26214511/)]
26. South DA, Skelley JW, Dang M, Woolley T. Near-miss transcription errors: a comparison of reporting rates between a novel error-reporting mechanism and a current formal reporting system. *Hosp Pharm* 2015 Feb;50(2):118-124 [FREE Full text] [doi: [10.1310/hpj5002-118](https://doi.org/10.1310/hpj5002-118)] [Medline: [25717207](https://pubmed.ncbi.nlm.nih.gov/25717207/)]
27. Elsaid K, Truong T, Monckeberg M, McCarthy H, Butera J, Collins C. Impact of electronic chemotherapy order forms on prescribing errors at an urban medical center: results from an interrupted time-series analysis. *Int J Qual Health Care* 2013 Dec;25(6):656-663. [doi: [10.1093/intqhc/mzt067](https://doi.org/10.1093/intqhc/mzt067)] [Medline: [24132956](https://pubmed.ncbi.nlm.nih.gov/24132956/)]
28. Champion TR, May AK, Waitman LR, Ozdas A, Gadd CS. Effects of blood glucose transcription mismatches on a computer-based intensive insulin therapy protocol. *Intensive Care Med* 2010 Sep;36(9):1566-1570 [FREE Full text] [doi: [10.1007/s00134-010-1868-7](https://doi.org/10.1007/s00134-010-1868-7)] [Medline: [20352190](https://pubmed.ncbi.nlm.nih.gov/20352190/)]

29. Ginsberg BH. Factors affecting blood glucose monitoring: sources of errors in measurement. *J Diabetes Sci Technol* 2009 Jul 01;3(4):903-913 [[FREE Full text](#)] [doi: [10.1177/193229680900300438](https://doi.org/10.1177/193229680900300438)] [Medline: [20144340](#)]
30. Prescrire Editorial Staff. Insulin use: preventable errors. *Prescrire Int* 2014 Jan;23(145):14-17. [Medline: [24516905](#)]

Abbreviations

EMR: electronic medical record
ICU: intensive care unit
POCT: point-of-care test
STICU: surgical trauma intensive care unit
VIN: visit identification number

Edited by C Lovis; submitted 08.08.18; peer-reviewed by K Glover, L Katz; comments to author 27.12.18; revised version received 10.01.19; accepted 11.01.19; published 25.03.19.

Please cite as:

Sowan AK, Vera A, Malshe A, Reed C

Transcription Errors of Blood Glucose Values and Insulin Errors in an Intensive Care Unit: Secondary Data Analysis Toward Electronic Medical Record-Glucometer Interoperability

JMIR Med Inform 2019;7(1):e11873

URL: <http://medinform.jmir.org/2019/1/e11873/>

doi: [10.2196/11873](https://doi.org/10.2196/11873)

PMID: [30907735](https://pubmed.ncbi.nlm.nih.gov/30907735/)

©Azizeh Khaled Sowan, Ana Vera, Ashwin Malshe, Charles Reed. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Extraction of Geriatric Syndromes From Electronic Health Record Clinical Notes: Assessment of Statistical Natural Language Processing Methods

Tao Chen¹, PhD; Mark Dredze², MA, PhD; Jonathan P Weiner³, DrPH; Leilani Hernandez⁴, MPH; Joe Kimura⁴, MPH, MD; Hadi Kharrazi^{3,5}, MD, PhD

¹Center for Language and Speech Processing, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States

²Department of Computer Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States

³Center for Population Health IT, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

⁴Academic Institute, Atrius Health, Boston, MA, United States

⁵Division of Health Sciences Informatics, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:

Tao Chen, PhD

Center for Language and Speech Processing

Whiting School of Engineering

Johns Hopkins University

Hackerman Hall 226

3400 North Charles Street

Baltimore, MD, 21218

United States

Phone: 1 410 516 4237

Fax: 1 410 516 5050

Email: tchen77@jhu.edu

Abstract

Background: Geriatric syndromes in older adults are associated with adverse outcomes. However, despite being reported in clinical notes, these syndromes are often poorly captured by diagnostic codes in the structured fields of electronic health records (EHRs) or administrative records.

Objective: We aim to automatically determine if a patient has any geriatric syndromes by mining the free text of associated EHR clinical notes. We assessed which statistical natural language processing (NLP) techniques are most effective.

Methods: We applied conditional random fields (CRFs), a widely used machine learning algorithm, to identify each of 10 geriatric syndrome constructs in a clinical note. We assessed three sets of features and attributes for CRF operations: a base set, enhanced token, and contextual features. We trained the CRF on 3901 manually annotated notes from 85 patients, tuned the CRF on a validation set of 50 patients, and evaluated it on 50 held-out test patients. These notes were from a group of US Medicare patients over 65 years of age enrolled in a Medicare Advantage Health Maintenance Organization and cared for by a large group practice in Massachusetts.

Results: A final feature set was formed through comprehensive feature ablation experiments. The final CRF model performed well at patient-level determination (macroaverage F1=0.834, microaverage F1=0.851); however, performance varied by construct. For example, at phrase-partial evaluation, the CRF model worked well on constructs such as absence of fecal control (F1=0.857) and vision impairment (F1=0.798) but poorly on malnutrition (F1=0.155), weight loss (F1=0.394), and severe urinary control issues (F1=0.532). Errors were primarily due to previously unobserved words (ie, out-of-vocabulary) and a lack of context.

Conclusions: This study shows that statistical NLP can be used to identify geriatric syndromes from EHR-extracted clinical notes. This creates new opportunities to identify patients with geriatric syndromes and study their health outcomes.

(*JMIR Med Inform* 2019;7(1):e13039) doi:[10.2196/13039](https://doi.org/10.2196/13039)

KEYWORDS

geriatrics; clinical notes; natural language processing; information extraction; conditional random fields

Introduction

Geriatric syndromes encompass multifactorial health conditions in older adults that generally do not fit into specific disease categories [1,2]. Geriatric syndromes, such as walking difficulty, falls, and incontinence, are often associated with adverse outcomes, such as morbidity, mortality, hospitalizations, and nursing home admissions [3]. Measuring geriatric syndromes at a population level is of great interest to health care providers and researchers to determine correlates of adverse outcomes. Hazra et al [4] contrasted health status, as measured by geriatric syndromes, of men and women aged 100 years or older. In addition, Anzaldi et al [5] measured prevalence of geriatric syndromes among individuals aged 65 years or older who had a mention of frailty in their clinical notes.

However, the multifactorial nature of geriatric syndromes complicates commonly accepted definitions for the recognition, diagnosis, and coding of these syndromes [1]. As a consequence, geriatric syndromes are often poorly captured by diagnostic codes (eg, International Classification of Diseases [ICD]9/10 codes) in the structured field of electronic health records (EHRs) [5], thus limiting research opportunities. Yet the unstructured clinical note (ie, free text) of EHRs contains rich information that describes geriatric syndromes. Considering the high adoption rate of EHRs among health care providers, unlocking information from clinical notes would yield a valuable resource for geriatric research and patient- and population-level interventions.

We propose a method to automatically identify patients exhibiting geriatric syndromes by analyzing text in clinical notes. We focus on 10 geriatric syndrome constructs: falls, malnutrition, dementia, severe urinary control issues, absence of fecal control, visual impairment, walking difficulty, pressure ulcers, lack of social support, and weight loss. We present methods based on natural language processing (NLP), specifically information extraction, that identify spans of text (ie, phrases) that attest to a geriatric syndrome. Previously, such clinical NLP has been leveraged to extract medical entities and concepts [6], such as disorder entity [7,8], medication [9], adverse drug reaction [10], smoking status [11], and risk [12].

Clinical NLP techniques can be roughly divided into two types: rule-based systems and machine learning (ie, statistical) techniques. Rule-based systems, which have long been the norm in the clinical setting, rely on manual definition of rules (eg, regular expressions) that identify phrases of interest in notes. For example, Anzaldi et al [5] and Kharrazi et al [13] developed a set of regular expressions to extract 10 geriatric syndrome constructs from clinical notes. The performance of rule-based approaches, however, requires significant domain expertise and is limited by the inflexibility of rule-based pattern matching. Therefore, statistical NLP methods based on machine learning techniques have long become the norm in the NLP community, with both academic research and industry systems relying almost exclusively on these methods [6]. Statistical methods rely on the construction of a manually annotated dataset to train a machine learning model; the trained model is then applied to extract entities and concepts from unannotated notes.

We propose to extract geriatric syndrome constructs from clinical notes using the conditional random field (CRF), a supervised statistical machine learning model that shows superior performance in many general and clinical information extraction tasks [3,7,8]. However, unlike most work in clinical NLP that focuses on well-defined entities and concepts (eg, based on Unified Medical Language System concepts), geriatric syndromes are often open-ended without clear definitions. For instance, the text spans of the geriatric syndrome constructs are much longer (ie, average length of 3.3 words) than well-defined entities and concepts (eg, average length of disorder entities are 1.8 tokens [ie, words]) [14]. Additionally, a lack of a clear coding standard for some of the geriatric syndrome constructs [1] means annotations are inconsistent, often including or excluding adjacent terms in the annotated construct mention. These challenges call into question the ability to develop a statistical NLP system for identifying patients with geriatric syndromes.

The objective of this paper is to explore the efficacy of a CRF model and various feature (ie, attribute) sets for the identification and classification of geriatric syndromes from clinical text. We consider the use of three feature sets: a base feature set, a token-enhanced set, and a set that includes contextual information. We evaluate the effectiveness of the model at identifying specific mentions (ie, phrases) as well as the overall ability to associate a construct with a patient (ie, aggregation over multiple notes). We report results on each of the 10 individual constructs and examine the factors that cause the accuracy of the trained model to vary over these constructs.

Methods

We begin with a description of the dataset followed by the clinical NLP model used to identify geriatric syndrome constructs. We then describe our data and experimental setting.

Dataset

We used anonymized EHR data shared with us by a large multispecialty medical group from New England, United States, for a cohort of elderly patients enrolled in a regional Medicare Advantage Health Maintenance Organization. We utilized a cohort of 18,341 members aged 65 years or older who received health insurance coverage between 2011 and 2013 and were assigned to this medical group as their primary source of medical care from this health plan. Clinical notes are documentations associated with patients' encounters; the top five encounter types were phone calls (37.8%), office visits (30.2%), refill (11.6%), letter (3.2%), and notation (2.6%). All data were stored on a secured network approved by the Institutional Review Board (IRB) of Johns Hopkins University School of Public Health (IRB number 6196).

For the analysis reported here, we randomly selected a small sample of 185 patients from the above cohort and included all of their unstructured clinical notes, resulting in a dataset of 8442 notes. These notes were manually annotated by three physicians to indicate mentions of the 10 constructs. All the physicians went through a training phase and coded a similar text to ensure an acceptable agreement could be reached before conducting

the annotation. Each note was annotated by one of the physicians. [Table 1](#) shows a few example annotations. The clinical notes were structured into sections indicated by a section header (eg, *Patient Medical History* and *Current Outpatient Prescriptions*). We segmented the notes by section using a list of keywords [15] and applied sentence segmentation and tokenization using the clinical Text Analysis Knowledge Extraction System (cTAKES) (Apache) [16], a clinical text-processing tool.

Clinical Natural Language Processing Algorithm

We modeled the construct identification and classification as a sequence labeling task. In a sequence labeling task, the model identifies and types spans of text according to established guidelines. Common sequence labeling tasks include part-of-speech taggings (ie, identifying each word as a noun, verb, or adjective) and named entity recognition (ie, identifying spans of text that refer to people, organizations, and locations by name). CRFs [17] are widely used statistical models for sequence labeling tasks in both traditional NLP [18] and clinical NLP [6-8]. In addition, CRFs were used by several validated systems in clinical note information extraction shared tasks [7,8]. We utilized the linear-chain CRF implementation from the CRFSuite software package developed by Naoaki Okazaki [19] for our work.

As a supervised machine learning algorithm, the CRF estimates (ie, learns) model parameters based on an annotated dataset (ie, training set). The trained model can then predict the labels of sequences without annotation. A key input to the model is a set of features: attributes of the input upon which the CRF builds a model and estimates parameters. Feature choices are a critical factor in determining the resulting performance of the model [17].

We designed and evaluated three sets of features to extract from the clinical note. These features capture basic information of the tokens (ie, words), enhanced information of the token, and the global context:

1. **Basic Features:** This set of features includes the lowercase of the token, the part-of-speech tag of the token, as identified by cTAKES, and three orthographic features that indicate whether the token is numeric, in uppercase, or in title case. In addition, we captured the local context of the token by incorporating these features from the previous and next token. In total, each token has 18 feature types.
2. **Enhanced Token Features:** This set of features captures additional features about the token. Many studies show that the stem of the token is a useful feature for information extraction tasks [20], thus we encoded it as another feature and explored how these features perform in the EHR domain. We used another two orthographic features to indicate whether the token is an ICD9 code [21] (hereafter,

Is-ICD9-Code; we could not evaluate ICD10 code, as our dataset predates the release of ICD10 code in 2015) and whether the token is a medical measurement unit (eg, kg or mL; hereafter, Is-Medical-Unit). To obtain these two features, we compared the token against the ICD9 codes [21] and common medical measurement units and encoded them as two binary features. The mention of a construct may not necessarily indicate the associated geriatric syndrome for the patient. The mention may be negated or reflect uncertainty in the diagnosis or could refer to another individual (eg, patient's family member). To identify such cases, we used cTAKES to obtain three attributes: negation, uncertainty, and subject of the construct entity mention. We encoded these as three features for each token (hereafter, Entity-Attributes).

3. **Global Context:** Our final feature set captures high-level context based on characteristics of the patient or clinical note. We first consider the section in which a token appears, as some section headers may suggest the mention does not refer to the patient (eg, *Family History*). Our second global context feature uses ICD9 codes mentioned in the text. Prior work [5,13] identified a list of 295 ICD9 codes that are indicative of the 10 constructs; [Table 2](#) details the number of ICD9 codes per construct. We leveraged this list to encode 10 binary features to indicate the mapping of a token to the 10 constructs (hereafter, ICD9-Annotation). That is, one binary feature that corresponds to the construct will be true when the token is in the list; otherwise, all 10 of the binary features are false. Alternatively, we utilized this list to modify the prediction from the CRF as a postprocessing step. When the token is present in the list, we altered its label to reflect the construct label from the list, even if the CRF failed to identify a corresponding construct mention (hereafter, ICD9-Annotation-Post).

Experimental Setting

Following standard evaluation conventions, we randomly split the 185 annotated patients into a training set (N=85), validation set (N=50), and test set (N=50). From [Table 3](#), we see the 10 constructs have skewed distributions. In the training set, the two most dominant constructs were walking difficulty and lack of social support, which were present in 66% (56/85) and 62% (53/85) of patients, respectively; the two rare constructs, malnutrition and pressure ulcers, were present in only 9% (8/85) and 11% (9/85) of patients, respectively. We estimated model parameters on the training set, tuned the hyperparameters of the training algorithm, chose features to include on the validation set, and evaluated our final trained model on the held-out test set. This evaluation procedure ensures that test set performance reflects real-world system accuracy, as choices of algorithm design and parameter estimates are made blind to the test set data.

Table 1. Example sentences from clinical notes that contain a construct: annotated construct phrases are italicized.

Construct	Example sentence from clinical notes (verbatim)
Absence of fecal control	She has also been experiencing urinary incontinence and a few <i>episodes of fecal incontinence</i> too.
Dementia	Patient <i>has dementia</i> and daughter feels as though it has worsened since Labor day.
Falls	She suffered <i>a fall</i> this past Tuesday and then was complaining of left shoulder pain.
Weight loss	Sed rate had been mildly elevated except the last one over 70 but in setting of acute illness and <i>weight loss</i> .
Malnutrition	<i>Inadequate energy intake</i> as evidenced by weight loss.
Pressure ulcers	She has <i>2 intragluteal decubitus</i> .
Lack of social support	She <i>is alone</i> at home much of the day.
Severe urinary control issues	She <i>has a suprapubic catheter</i> in (placed under interventional radiology at) because she was having pain on urination.
Visual impairment	Has been seen by vision rehab and <i>is registered with of blind</i> .
Walking difficulty	<i>Ambulates slowly</i> , uses vital signs as above.

Table 2. Statistics related to the 10 constructs.

Construct	Number of ICD9 ^a codes that indicate a construct (n)	Average number of tokens per construct	Average number of mentions per patient in the test set	Perplexity on test set ^b
Absence of fecal control	2	2.98	2.67	11.30
Dementia	58	2.76	13.00	26.28
Falls	45	3.37	9.04	57.68
Weight loss	15	3.01	13.53	33.80
Malnutrition	26	2.04	13.92	100.64
Pressure ulcers	35	3.48	5.67	66.90
Lack of social support	14	4.03	15.23	29.96
Severe urinary control issues	14	2.94	13.71	117.48
Visual impairment	55	3.62	9.31	57.68
Walking difficulty	31	3.43	12.59	84.27

^aICD9: International Classification of Diseases 9.

^bPerplexity is computed on the test set based on the construct-specific language model trained on the training set: detailed in the Error Analysis section.

We performed feature ablation experiments on the validation set to assess the effectiveness of the proposed features. In each experiment, we trained the CRF on the training set using basic features and one or more features from the proposed feature sets, evaluating the trained model on the validation set. For each experiment, we used the validation set to tune model hyperparameters in a grid search manner. Example hyperparameters are L2 regularizer and maximum number of iterations; the learning rate of stochastic gradient descent (SGD) is automatically determined by CRFSuite. We also assessed both the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm and SGD optimization methods, selecting SGD as our final training algorithm for its better performance. Finally, we trained the model with the best feature combination and hyperparameter settings and reported the final performance on the test set.

We consider four different evaluation metrics for our CRF. The most restrictive evaluation measure is phrase-exact, in which

we mark a prediction as correct only if the extracted and identified phrase exactly matched the labeled phrase. Under this metric, for the sentence “This patient walks with a walker,” an answer would only be correct if it identified the phrase “with a walker” (ie, the provided annotation) as walking difficulty but not if it selected “walks with a walker.” While this type of evaluation is standard in many information extraction tasks, it is too strict considering that our goal is to associate a geriatric syndrome with the patient. On the other hand, we also observed our three annotators sometimes exhibiting inconsistency in including or excluding unimportant words (eg, prepositions and verbs) in the annotated phrase (eg, “with a walker,” “walks with a walker,” or “walker”). Therefore, we consider a partial matching metric (ie, phrase-partial), which marks a prediction as correct if the predicted phrase overlaps with the manual annotation. In this setting, the above prediction for walking difficulty would be correct. Such partial matching has also been adopted in other clinical information extraction tasks [7,8].

Table 3. The construct and nonconstruct distribution among three datasets based on manual annotation.

Construct	Training set (3901 notes)		Validation set (1739 notes)		Test set (2802 notes)	
	Token ^a (N=1,083,670), n (%)	Patient ^b (N=85), n (%)	Token (N=435,851), n (%)	Patient (N=50), n (%)	Token (N=638,694), n (%)	Patient (N=50), n (%)
Absence of fecal control	126 (0.01)	12 (14)	126 (0.03)	4 (8)	34 (0.01)	3 (6)
Dementia	631 (0.06)	15 (18)	276 (0.06)	9 (18)	403 (0.06)	10 (20)
Falls	1419 (0.13)	37 (44)	293 (0.07)	21 (42)	748 (0.12)	23 (46)
Weight loss	365 (0.03)	21 (25)	263 (0.06)	14 (28)	752 (0.12)	19 (38)
Malnutrition	115 (0.01)	8 (9)	82 (0.02)	5 (10)	312 (0.05)	12 (24)
Pressure ulcers	308 (0.03)	9 (11)	18 (0.00)	4 (8)	126 (0.02)	6 (12)
Lack of social support	2026 (0.19)	53 (62)	1410 (0.32)	30 (60)	1691 (0.26)	30 (60)
Severe urinary control issues	694 (0.06)	16 (19)	81 (0.02)	4 (8)	323 (0.05)	7 (14)
Visual impairment	324 (0.03)	16 (19)	141 (0.03)	6 (12)	395 (0.06)	13 (26)
Walking difficulty	2253 (0.21)	56 (66)	1315 (0.30)	26 (52)	1423 (0.22)	34 (68)
Nonconstruct	1,075,409 (99.24)	85 (100)	431,846 (99.08)	50 (100)	632,487 (99.03)	50 (100)

^aDenotes the number of tokens in the dataset that were labeled as certain constructs.

^bDenotes the number of patients in the dataset who were identified containing certain constructs.

Since our goal was to associate constructs with patients, we considered two additional metrics. First, we identified the prediction of a construct as correct if that construct appeared anywhere in the clinical note (ie, note-level), which may occur when an annotator missed a construct mention. We also considered a patient-level evaluation, in which we marked a prediction as correct if the patient had the associated construct (eg, an annotator identified it somewhere in an associated note).

We computed the popular information extraction metrics of precision (ie, true positive/[true positive + false positive]), recall (ie, true positive/[true positive + false negative]), and their harmonic mean (F1 or F measure). These metrics are related to sensitivity (ie, true positive rate) and specificity (ie, true negative rate). We provide these metrics for each of the four evaluation types and report model performance on each of the 10 constructs. We also report micro- and macroaveraged results, where microaveraging computes the average over every construct mention and macroaveraging gives equal weight to every construct. The difference reflects the variations of prevalence among the constructs. For all validation set choices, we used microaveraged F1 on the phrase-partial matching metric.

Error Analysis

To gain more insights, we performed an in-depth analysis on the system's errors. We quantified the chances that the CRF model would confuse the mention of one construct with another. Additionally, we trained 10 construct-specific bigram language models (ie, a probability distribution to estimate the relative likelihood of text) using the construct's mention texts from the

training set. We then computed the perplexity (ie, a measurement of how well a probability distribution predicts a sample) of each construct language model on mentions of the construct in the test set. In short, the perplexity captures how "surprised" the model would be by a construct reference in the test data based on how the construct was referenced in the training data.

Ethical Considerations

This study was approved by the IRB of Johns Hopkins University School of Public Health (IRB number 6196). Participant consent was not required as data was deidentified prior to analysis.

Results

We measured the micro- and macroaverage phrase-partial results of models with optimal hyperparameters on the validation data (see Table 4). Using the basic feature set alone, the CRF achieved a macroaverage F1 score of 0.583 and a microaverage F1 score of 0.727. Most additional features improved the F1 score. Of these, the most effective single feature was the stem, which improved the macroaverage F1 score by 0.103, a relative improvement of 17.7%, and improved the microaverage F1 score by 0.033, a relative improvement of 4.5%, when compared to the basic model using basic features. The two exceptions were token-enhanced features: Is-ICD9-Code and Is-Medical-Unit. Though the two features did not improve the F1 score, they did increase the microaverage precision from 0.93 to 0.959 (Is-ICD9-Code) and 0.948 (Is-Medical-Unit). This suggests that they could still contribute to overall improvements when combined with other features focused on improving recall.

Table 4. Phrase-partial evaluation on the validation set.

Feature set and features	<i>P</i> value ^a	Macroaverage			Microaverage		
		Precision	Recall	F1	Precision	Recall	F1
Basic features							
Basic	N/A ^b	0.828	0.450	0.583	0.930	0.597	0.727
Enhanced token features							
B ^c +Is-ICD9 ^d -Code	<.001	0.874	0.472	0.613	0.887	0.640	0.744
B+Is-Medical-Unit	<.001	0.828	0.402	0.541	0.959	0.538	0.689
B+Entity-Attributes	<.001	0.823	0.398	0.537	0.948	0.528	0.678
B+Stem	.03	0.856	0.572	0.686	0.864	0.678	0.760
Contextual features							
B+Section	<.001	0.783	0.544	0.642	0.874	0.682	0.766
B+ICD9-Annotation	<.001	0.888	0.462	0.608	0.928	0.598	0.727
B+ICD9-Annotation-Post	<.001	0.823	0.478	0.605	0.912	0.604	0.727
Combination (B+Enhanced+Context)							
B+all Enhanced (C ^e +U ^f +E ^g +S ^h)+all Context (T ⁱ +A ^j +AP ^k)	<.001	0.793	0.633	0.704	0.757	0.714	0.735
B+Enhanced (C+E+S)+all Context (T+A+AP)	<.001	0.837	0.483	0.613	0.895	0.546	0.678
B+Enhanced (C+E+S)+Context (A+AP)	<.001	0.874	0.529	0.659	0.906	0.630	0.743
<i>B+all Enhanced (C+U+E+S)+Context (A+AP)</i> ^l	<.001	0.862	0.567	0.684	0.880	0.681	0.768
B+Enhanced (C+S)+Context (A+AP)	<.001	0.799	0.509	0.622	0.896	0.616	0.730
Non-CRF^m model							
Only uses annotated ICD9 codes as a rule to identify constructs	<.001	0.803	0.139	0.236	0.885	0.059	0.111

^aWe conducted McNemar's test to measure the difference between the results of using basic features and other features.

^bN/A: not applicable.

^cB: basic.

^dICD9: International Classification of Diseases 9.

^eC: Is-ICD9-Code.

^fU: Is-Medical-Unit.

^gE: Entity-Attributes.

^hS: stem.

ⁱT: section.

^jA: ICD9-Annotation.

^kAP: ICD9-Annotation-Post.

^lThe best-performing model is italicized.

^mCRF: conditional random field.

In contrast, the section feature leads to a significant improvement on recall but lowers the precision by a large margin. Therefore, we further evaluated combinations of these three features and other features on the validation set. The best-performing model (L2 regularizer=0.2, maximum number of iterations=100) on validation data was trained with all the features except section and achieved a macroaverage F1 score of 0.684 and a microaverage F1 score of 0.768. This model significantly ($P<.001$ via McNemar's test) outperformed the basic model with a macroaverage F1 score of 0.101, a relative improvement of 17.3%, and a microaverage F1 score of 0.041,

a relative improvement of 8.8%. Additionally, we implemented a simple rule-based method that only uses the annotated ICD9 codes to identify constructs in a sentence. As expected, this method achieved a very low recall (0.139 at macroaverage and 0.059 at microaverage) and F1 score (0.236 at macroaverage and 0.111 at microaverage), which further validates that geriatric syndromes are poorly captured by diagnosis code. This also demonstrates the importance of developing models to identify constructs by mining unstructured text.

We evaluated this best-performing CRF model on the test set and report per-construct results in addition to overall averages

(see Table 5 and Figure 1). The CRF obtained macroaverage F1 scores of 0.394 for phrase-exact, 0.666 for phrase-partial, 0.759 for note, and 0.834 for patient. Microaverage F1 scores were 0.410 for phrase-exact, 0.661 for phrase-partial, 0.787 for note, and 0.851 for patient. Across all constructs, precision was higher than recall, meaning that the model favored accurate predictions over coverage of construct mentions. By relying on this high-precision approach and the repetition of construct mentions in patient's clinical notes (see Table 2) (eg, the lack of social support was mentioned, on average, 15.23 times per patient), our model obtained a much higher performance on patient than on the other three evaluations—phrase-exact, phrase-partial, and note.

Performance varied widely for different constructs (see Table 5 and Figure 1). On a phrase-partial analysis, seven constructs generated an F1 score of over 0.7, of which absence of fecal control was the best (0.857), while the three worst constructs—malnutrition, weight loss, and severe urinary control issues—obtained scores of 0.155, 0.394, and 0.532, respectively. To understand the F1 variance across the constructs, we performed an in-depth error analysis of system errors. Overall, 97.0% of errors were caused by missing the constructs; the model seldomly confused mentions of one construct for another.

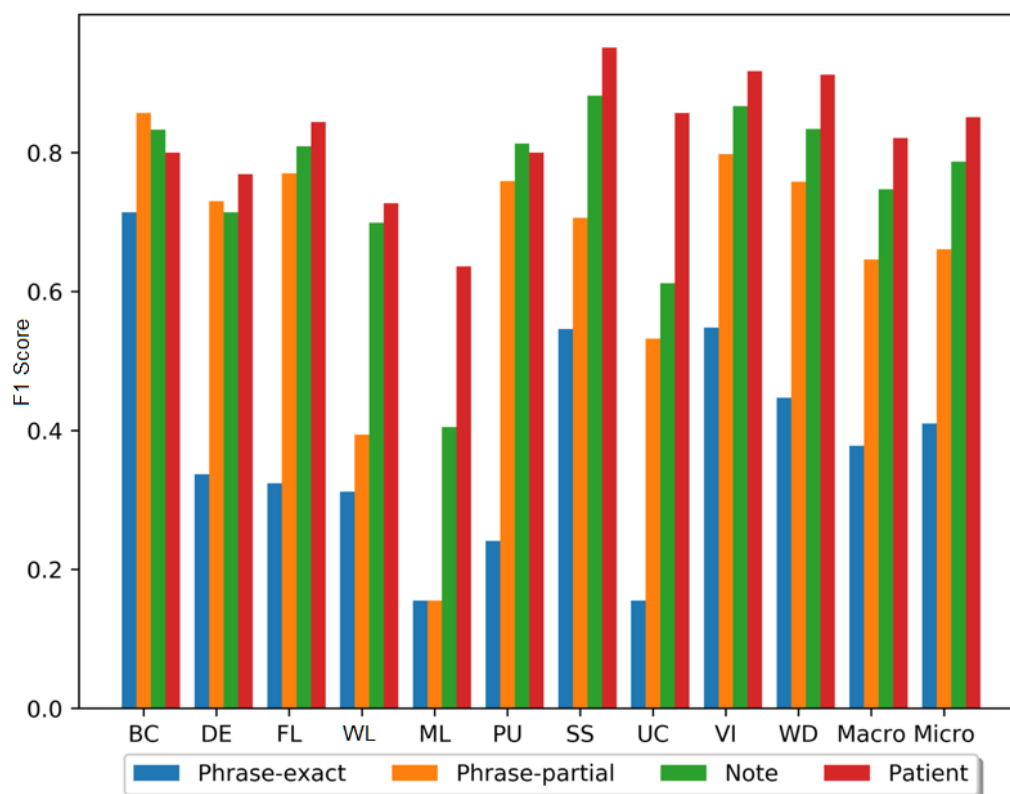
One primary cause of low recall was the limited training instances, which limited the variations observed during training. The training set contains 3901 notes from 85 patients, but only 0.76% of tokens indicated any of the 10 geriatric syndromes. For the three poorly performing constructs, we found that both malnutrition and severe urinary control issues had a high out-of-vocabulary (OOV) rate in the test set (ie, many of the words used to refer to these constructs were unobserved during training).

We measured the perplexity scores for each construct (see Table 2). This score measures how unprepared the model would be by the construct's mention in the test set. Malnutrition (100.64) and severe urinary control issues (117.48) showed a much higher perplexity score than the other constructs, confirming higher OOV rates as compared to the other constructs. However, the poor performance of weight loss was primarily caused by confusion between intentional weight loss (ie, nonconstruct; patients who are overweight and thus trying to lose weight) and unintentional weight loss (ie, our construct). Since intentional weight loss is dominant, the CRF identified incidents of weight loss as nonconstruct, yielding a very low recall (0.272) and F1 score (0.394).

Table 5. The evaluation results of the best-performing model on the test set.

Construct	Phrase-exact			Phrase-partial			Note			Patient		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Absence of fecal control	0.833	0.625	0.714	1	0.750	0.857	1	0.714	0.833	1	0.667	0.800
Dementia	0.324	0.350	0.337	0.703	0.759	0.730	0.604	0.873	0.714	0.625	1	0.769
Falls	0.387	0.279	0.324	0.942	0.651	0.770	0.926	0.719	0.809	0.864	0.826	0.844
Weight loss	0.571	0.215	0.312	0.714	0.272	0.394	0.866	0.586	0.699	0.857	0.632	0.727
Malnutrition	0.577	0.090	0.155	0.577	0.090	0.155	0.680	0.288	0.405	0.700	0.583	0.636
Pressure ulcers	0.304	0.200	0.241	0.957	0.629	0.759	0.929	0.722	0.813	1	0.667	0.800
Lack of social support	0.551	0.541	0.546	0.707	0.706	0.706	0.923	0.845	0.882	0.935	0.967	0.951
Severe urinary control issues	0.207	0.124	0.155	0.690	0.433	0.532	0.682	0.556	0.612	0.857	0.857	0.857
Visual impairment	0.687	0.456	0.548	1	0.664	0.798	1	0.765	0.867	1	0.846	0.917
Walking difficulty	0.517	0.394	0.447	0.842	0.689	0.758	0.894	0.781	0.834	0.912	0.912	0.912
Macroaverage	0.496	0.327	0.394	0.813	0.564	0.666	0.850	0.685	0.759	0.875	0.796	0.834
Microaverage	0.493	0.351	0.410	0.785	0.571	0.661	0.806	0.726	0.787	0.868	0.834	0.851

Figure 1. The F1 scores of the best-performing model on the test set. BC: absence of fecal control; DE: dementia; FL: fall; WL: weight loss; ML: malnutrition; PU: pressure ulcers; SS: lack of social support; UC: severe urinary control issues; VI: visual impairment; WD: walking difficulty; Macro: macroaverage; Micro: microaverage.



Discussion

Principal Findings

Geriatric syndromes are often not captured in EHRs' structured fields, which requires examining of EHRs' free text to identify older adults with such syndromes [5,13]. We applied NLP techniques to identify patients who have geriatric syndromes by analyzing their clinical notes. We trained CRF models with three sets of features to identify phrases that indicate geriatric syndromes and aggregated identified phrases to make patient-level determinations. Our best-performing model obtained a macroaverage F1 score of 0.834 and microaverage F1 score of 0.851 for identifying 10 geriatric syndrome constructs for patients. Our system identified most patients with geriatric syndromes by mining EHRs' clinical notes and could support research on geriatric syndromes.

Technical Challenges and Opportunities

Extracting and identifying geriatric syndromes is much more difficult than well-defined entity extraction. For instance, in the disorder entity extraction task of the 2013 Shared Annotated Resources/Conference and Labs of the Evaluation Forum (ShARE/CLEF) eHealth Evaluation Lab [22], a shared task focusing on clinical NLP, the best-performing system [23] also relied on CRF models but obtained much better results (phrase-partial F1=0.873) than did ours (macroaverage F1=0.666, microaverage F1=0.661). Our model heavily favors precision over recall (eg, macroaverage precision=0.813 vs macroaverage recall=0.564 at phrase-partial) and exhibits varied performance on different constructs; it works well on the

absence of fecal control (F1=0.857), visual impairment (F1=0.798), and fall (F1=0.770), but poorly on malnutrition (F1=0.155), weight loss (F1=0.394), and severe urinary control issues (F1=0.532). These variations are primarily caused by the high OOV rates of malnutrition and severe urinary control issues as well as the confusion between intentional weight loss (ie, nonconstruct) and unintentional weight loss (ie, our construct). These challenges may limit statistical NLP models in geriatric syndrome research since the models will miss a large portion of patients with malnutrition, severe urinary control issues, and weight loss.

Our error analysis suggests several directions of future work. We can lower OOV rates by obtaining additional annotated clinical notes. However, manual annotation is time-consuming and costly (eg, it took approximately 240 person-hours in total for the three experts to annotate 185 patients); we thus prefer other solutions. One approach would be to generalize representations away from lexical forms, replacing words or phrases with embeddings. Embeddings are a form of high-dimensional dense vector representation where similar words or phrases tend to be close to each other. Such embeddings could be trained with a large number of unlabeled notes in an unsupervised fashion [24,25], which have shown to be effective in clinical NLP tasks [26,27]. Since we can learn embeddings even for words unseen in training, this may mitigate the OOV problem. In addition, we should identify the larger context of the mention to differentiate intentional and unintentional weight loss. For example, if the note discusses obesity, it is a strong indicator that the weight loss is intentional. This contextual information may be reflected in the diagnostic

codes in the structured field of EHRs or the larger context in the free text. We could incorporate the contextual information into the extraction model through the use of learned representations of the context (ie, embeddings). Finally, recent interest in deep learning model architectures [28] have shown promise on clinical data [29,30]. These models may provide added benefits over CRFs.

Clinical and Population Health Implications

Our work has implications for managing older adult patients by enabling clinicians and researchers to identify a broader set of patients with geriatric syndromes using EHRs' free text. For example, a wider identification of geriatric syndromes enables clinicians to adjust interventions and provides researchers the opportunity to expand study cohorts [13]. However, extracting geriatric syndromes from clinical notes may require dealing with multiple EHR issues, such as lack of data-quality specifications for EHRs and increased rate of missing data over time [31], challenges with incorporating questionnaires within the EHRs' architecture (eg, geriatric frailty questionnaires) [32], and variation of EHR use and maturation across different health delivery systems [33].

Population health management efforts are increasingly using EHR data for risk stratification of patient populations [34-38]. Our model increases the coverage of risk stratification models developed specifically for an older adult population. Identification of new cases of geriatric syndromes will enable care coordinators and case managers to better manage patient populations [34], which can lead to more streamlined and efficient workflows. Furthermore, our work has implications for extracting nonclinical information, such as social determinants of health (SDH)—an open-ended construct similar to geriatric syndromes—from EHRs' free text. SDH is a combination of lifestyle, behavioral, social, economic, and environmental factors that are powerful drivers of health and well-being [39]. Similar to geriatric syndromes, SDH are poorly captured by diagnostic codes in EHRs [40-42]. More broadly, our success with NLP techniques for geriatric syndromes suggests that other areas that actively use SDH information to bridge the gap between population and public health may similarly benefit from EHRs' free text and NLP techniques [43-46].

Limitations

This work has several limitations. First, given the evolving concept of geriatric syndromes and their varied contextual information mentioned in clinical notes, and despite the rigorous training of the annotators, annotations were slightly inconsistent in including or excluding unrelated contextual wording (eg, including or excluding the location of a fall in addition to the event of a fall). Additionally, from error analysis, we also observed a few cases where our CRF model correctly identified the construct mentions but were missed by human annotators. These issues could be alleviated when soliciting multiple experts to repeatedly annotate the notes. Second, we only experimented with the 185 annotated patients, while our dataset contains a large portion of unlabeled notes (ie, 18,156 patients). It will be interesting to apply our CRF model to the unlabeled notes and see how these 10 geriatric syndromes distribute in the larger population. Third, we limited our study to 10 constructs of geriatric syndromes, but many other types of geriatric syndromes (eg, delirium and functional decline) exist. We plan to generalize our model to extract other geriatric syndromes when their annotations are available. Finally, we did not analyze the temporal patterns of the geriatric syndrome constructs. Future studies should investigate whether the temporal patterns of a construct, especially if deemed resolvable over time such as lack of social support, will be altered differently by different NLP solutions (eg, measuring temporal accuracy).

Conclusions

We have presented an NLP solution for the automatic identification and classification of patients exhibiting geriatric syndromes by analyzing free text in patient clinical notes. We formulated the problem as an information extraction task and trained a CRF model to identify geriatric syndrome constructs from text. We presented enhanced features for this task and created a final system that obtains a microaverage F1 of 0.851 for patient-level determination of constructs. Our error analysis revealed that new words and a lack of context account for the worst-performing constructs. Future work should develop strategies that do not require additional training annotations to mitigate these errors.

Acknowledgments

This work was supported by Atrius Health and the Center for Population Health IT, Johns Hopkins University.

Conflicts of Interest

None declared.

References

1. Inouye SK, Studenski S, Tinetti ME, Kuchel GA. Geriatric syndromes: Clinical, research, and policy implications of a core geriatric concept. *J Am Geriatr Soc* 2007 May;55(5):780-791 [FREE Full text] [doi: [10.1111/j.1532-5415.2007.01156.x](https://doi.org/10.1111/j.1532-5415.2007.01156.x)] [Medline: [17493201](https://pubmed.ncbi.nlm.nih.gov/17493201/)]
2. Olde Rikkert MGM, Rigaud AS, van Hoeyweghen RJ, de Graaf J. Geriatric syndromes: Medical misnomer or progress in geriatrics? *Neth J Med* 2003 Mar;61(3):83-87 [FREE Full text] [Medline: [12765229](https://pubmed.ncbi.nlm.nih.gov/12765229/)]

3. Wang S, Shamliyan TA, Talley KMC, Ramakrishnan R, Kane RL. Not just specific diseases: Systematic review of the association of geriatric syndromes with hospitalization or nursing home admission. *Arch Gerontol Geriatr* 2013;57(1):16-26. [doi: [10.1016/j.archger.2013.03.007](https://doi.org/10.1016/j.archger.2013.03.007)] [Medline: [23578847](https://pubmed.ncbi.nlm.nih.gov/23578847/)]
4. Hazra N, Dregan A, Jackson S, Gulliford M. Differences in health at age 100 according to sex: Population-based cohort study of centenarians using electronic health records. *J Am Geriatr Soc* 2015 Jul;63(7):1331-1337 [FREE Full text] [doi: [10.1111/jgs.13484](https://doi.org/10.1111/jgs.13484)] [Medline: [26096699](https://pubmed.ncbi.nlm.nih.gov/26096699/)]
5. Anzaldi LJ, Davison A, Boyd CM, Leff B, Kharrazi H. Comparing clinician descriptions of frailty and geriatric syndromes using electronic health records: A retrospective cohort study. *BMC Geriatr* 2017 Dec 25;17(1):248 [FREE Full text] [doi: [10.1186/s12877-017-0645-7](https://doi.org/10.1186/s12877-017-0645-7)] [Medline: [29070036](https://pubmed.ncbi.nlm.nih.gov/29070036/)]
6. Wang Y, Wang L, Rastegar-Mojarad M, Moon S, Shen F, Afzal N, et al. Clinical information extraction applications: A literature review. *J Biomed Inform* 2018 Dec;77:34-49 [FREE Full text] [doi: [10.1016/j.jbi.2017.11.011](https://doi.org/10.1016/j.jbi.2017.11.011)] [Medline: [29162496](https://pubmed.ncbi.nlm.nih.gov/29162496/)]
7. Elhadad N, Pradhan S, Gorman S, Manandhar S, Chapman W, Savova G. SemEval-2015 task 14: Analysis of clinical text. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). 2015 Presented at: 9th International Workshop on Semantic Evaluation (SemEval 2015); June 4-5, 2015; Denver, CO p. 303-310.
8. Pradhan S, Elhadad N, Chapman W, Manandhar S, Savova G. Semeval-2014 task 7: Analysis of clinical text. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). 2014 Presented at: 8th International Workshop on Semantic Evaluation (SemEval 2014); August 23-24, 2014; Dublin, Ireland p. 54-62.
9. Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: A medication information extraction system for clinical narratives. *J Am Med Inform Assoc* 2010;17(1):19-24 [FREE Full text] [doi: [10.1197/jamia.M3378](https://doi.org/10.1197/jamia.M3378)] [Medline: [20064797](https://pubmed.ncbi.nlm.nih.gov/20064797/)]
10. Iqbal E, Mallah R, Jackson RG, Ball M, Ibrahim ZM, Broadbent M, et al. Identification of adverse drug events from free-text electronic patient records and information in a large mental health case register. *PLoS One* 2015;10(8):e0134208 [FREE Full text] [doi: [10.1371/journal.pone.0134208](https://doi.org/10.1371/journal.pone.0134208)] [Medline: [26273830](https://pubmed.ncbi.nlm.nih.gov/26273830/)]
11. Sohn S, Savova GK. Mayo clinic smoking status classification system: Extensions and improvements. *AMIA Annu Symp Proc* 2009 Nov 14;2009:619-623 [FREE Full text] [Medline: [20351929](https://pubmed.ncbi.nlm.nih.gov/20351929/)]
12. Deleger L, Brodzinski H, Zhai H, Li Q, Lingren T, Kirkendall ES, et al. Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. *J Am Med Inform Assoc* 2013 Dec;20(e2):e212-e220 [FREE Full text] [doi: [10.1136/amiajnl-2013-001962](https://doi.org/10.1136/amiajnl-2013-001962)] [Medline: [24130231](https://pubmed.ncbi.nlm.nih.gov/24130231/)]
13. Kharrazi H, Anzaldi LJ, Hernandez L, Davison A, Boyd CM, Leff B, et al. The value of unstructured electronic health record data in geriatric syndrome case identification. *J Am Geriatr Soc* 2018 Aug;66(8):1499-1507. [doi: [10.1111/jgs.15411](https://doi.org/10.1111/jgs.15411)] [Medline: [29972595](https://pubmed.ncbi.nlm.nih.gov/29972595/)]
14. Kelly L, Goeriot L, Suominen H, Schreck T, Leroy G, Mowery D, et al. Overview of the ShARe/CLEF eHealth Evaluation Lab 2014. In: Proceedings of the 5th International Conference of the CLEF Initiative, CLEF 2014. Cham, Switzerland: Springer International Publishing; 2014 Presented at: 5th International Conference of the CLEF Initiative, CLEF 2014; September 15-18, 2014; Sheffield, UK p. 172-191. [doi: [10.1007/978-3-319-11382-1_17](https://doi.org/10.1007/978-3-319-11382-1_17)]
15. GitHub. Section headers used by cTAKES URL: https://github.com/healthnlp/examples/blob/master/ctakes-web-client/src/main/resources/org/apache/ctakes/core/sections/ccda_sections.txt [accessed 2019-03-18] [WebCite Cache ID [76Gv8ijKy](https://www.webcitation.org/76Gv8ijKy)]
16. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *J Am Med Inform Assoc* 2010;17(5):507-513 [FREE Full text] [doi: [10.1136/jamia.2009.001560](https://doi.org/10.1136/jamia.2009.001560)] [Medline: [20819853](https://pubmed.ncbi.nlm.nih.gov/20819853/)]
17. Lafferty J, McCallum A, Pereira F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proceedings of the Eighteenth International Conference on Machine Learning. 2001 Presented at: Eighteenth International Conference on Machine Learning; June 28-July 01, 2001; Williamstown, MA p. 282-289.
18. Getoor L, Taskar B, editors. Introduction to Statistical Relational Learning. Cambridge, MA: MIT Press; 2007.
19. Okazaki N. Chokkan.org. CRFsuite: A fast implementation of Conditional Random Fields (CRFs) URL: <http://www.chokkan.org/software/crfsuite> [accessed 2018-12-02] [WebCite Cache ID [74N6saDQb](https://www.webcitation.org/74N6saDQb)]
20. Nadeau D, Sekine S. A survey of named entity recognition and classification. *Int J Linguist Lang Resour* 2007;30(1):3-26. [doi: [10.1075/li.30.1.03nad](https://doi.org/10.1075/li.30.1.03nad)]
21. CMS.gov. ICD-9-CM diagnosis and procedure codes: Abbreviated and full code titles URL: <https://www.cms.gov/medicare/coding/icd9providerdiagnosticcodes/codes.html> [accessed 2019-03-15] [WebCite Cache ID [76spyCXVX](https://www.webcitation.org/76spyCXVX)]
22. Suominen H, Salanterä S, Velupillai S, Chapman W, Savova G, Elhadad N, et al. Overview of the ShARe/CLEF eHealth Evaluation Lab 2013. In: Proceedings of the 4th International Conference of the CLEF Initiative, CLEF 2013. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 Presented at: 4th International Conference of the CLEF Initiative, CLEF 2013; September 23-26, 2013; Valencia, Spain p. 212-231. [doi: [10.1007/978-3-642-40802-1_24](https://doi.org/10.1007/978-3-642-40802-1_24)]
23. Tang B, Wu Y, Jiang M, Denny J, Xu H. Recognizing and encoding disorder concepts in clinical text using machine learning and vector space model. In: Workshop Proceedings of the 4th International Conference of the CLEF Initiative: Working Notes for CLEF 2013 Conference. 2013 Presented at: 4th International Conference of the CLEF Initiative: Working Notes for CLEF 2013 Conference; September 23-26, 2013; Valencia, Spain.

24. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013 Presented at: 26th International Conference on Neural Information Processing Systems; December 5-10, 2013; Lake Tahoe, NV p. 3111-3119.
25. Pennington J, Socher R, Manning C. Glove: Global vectors for word representation. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2014 Presented at: Conference on Empirical Methods in Natural Language Processing; October 25-29, 2014; Doha, Qatar p. 1532-1543.
26. Wu Y, Jiang M, Lei J, Xu H. Named entity recognition in Chinese clinical text using deep neural network. *Stud Health Technol Inform* 2015;216:624-628 [FREE Full text] [Medline: 26262126]
27. Liu Y, Ge T, Mathews K, Ji H, McGuinness D. Exploiting task-oriented resources to learn word embeddings for clinical abbreviation expansion. In: Proceedings of the Workshop on Biomedical Natural Language Processing (BioNLP 2015). 2015 Presented at: Workshop on Biomedical Natural Language Processing (BioNLP 2015); July 30, 2015; Beijing, China p. 92-97. [doi: [10.18653/v1/W15-3810](https://doi.org/10.18653/v1/W15-3810)]
28. Huang Z, Xu W, Yu K. arXiv. 2015. Bidirectional LSTM-CRF models for sequence tagging URL: <https://arxiv.org/pdf/1508.01991> [accessed 2019-03-18] [WebCite Cache ID 776BXemQ5]
29. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J Biomed Health Inform* 2018 Dec;22(5):1589-1604. [doi: [10.1109/JBHI.2017.2767063](https://doi.org/10.1109/JBHI.2017.2767063)] [Medline: 29989977]
30. Jagannatha AN, Yu H. Bidirectional RNN for medical event detection in electronic health records. *Proc Conf* 2016 Jun;2016:473-482 [FREE Full text] [Medline: 27885364]
31. Kharrazi H, Wang C, Scharfstein D. Prospective EHR-based clinical trials: The challenge of missing data. *J Gen Intern Med* 2014 Jul;29(7):976-978 [FREE Full text] [doi: [10.1007/s11606-014-2883-0](https://doi.org/10.1007/s11606-014-2883-0)] [Medline: 24839057]
32. Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice: Adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *J Clin Epidemiol* 2013 Aug;66(8 Suppl):S12-S20 [FREE Full text] [doi: [10.1016/j.jclinepi.2013.04.005](https://doi.org/10.1016/j.jclinepi.2013.04.005)] [Medline: 23849145]
33. Kharrazi H, Gonzalez C, Lowe K, Huerta T, Ford E. Forecasting the maturation of electronic health record functions among US hospitals: Retrospective analysis and predictive model. *J Med Internet Res* 2018 Aug 07;20(8):e10458 [FREE Full text] [doi: [10.2196/10458](https://doi.org/10.2196/10458)] [Medline: 30087090]
34. Kan H, Kharrazi H, Leff B, Boyd C, Davison A, Chang H, et al. Defining and assessing geriatric risk factors and associated health care utilization among older adults using claims and electronic health records. *Med Care* 2018 Dec;56(3):233-239. [doi: [10.1097/MLR.0000000000000865](https://doi.org/10.1097/MLR.0000000000000865)] [Medline: 29438193]
35. Kharrazi H, Chi W, Chang HY, Richards TM, Gallagher JM, Knudson SM, et al. Comparing population-based risk-stratification model performance using demographic, diagnosis and medication data extracted from outpatient electronic health records versus administrative claims. *Med Care* 2017 Dec;55(8):789-796. [doi: [10.1097/MLR.0000000000000754](https://doi.org/10.1097/MLR.0000000000000754)] [Medline: 28598890]
36. Kharrazi H, Weiner J. A practical comparison between the predictive power of population-based risk stratification models using data from electronic health records versus administrative claims: Setting a baseline for future EHR-derived risk stratification models. *Med Care* 2018 Dec;56(2):202-203. [doi: [10.1097/MLR.0000000000000849](https://doi.org/10.1097/MLR.0000000000000849)] [Medline: 29200132]
37. Chang HY, Richards TM, Shermock KM, Elder Dalpoas S, J Kan H, Alexander GC, et al. Evaluating the impact of prescription fill rates on risk stratification model performance. *Med Care* 2017 Dec;55(12):1052-1060. [doi: [10.1097/MLR.0000000000000825](https://doi.org/10.1097/MLR.0000000000000825)] [Medline: 29036011]
38. Lemke K, Gudzone K, Kharrazi H, Weiner J. Assessing markers from ambulatory laboratory tests for predicting high-risk patients. *Am J Manag Care* 2018 Jun 01;24(6):e190-e195 [FREE Full text] [Medline: 29939509]
39. Bazemore A, Cottrell E, Gold R, Hughes L, Phillips R, Angier H, et al. "Community vital signs": Incorporating geocoded social determinants into electronic records to promote patient and population health. *J Am Med Inform Assoc* 2016 Mar;23(2):407-412. [doi: [10.1093/jamia/ocv088](https://doi.org/10.1093/jamia/ocv088)] [Medline: 26174867]
40. Torres J, Lawlor J, Colvin J, Sills M, Bettenhausen J, Davidson A, et al. ICD social codes: An underutilized resource for tracking social needs. *Med Care* 2017 Dec;55(9):810-816. [doi: [10.1097/MLR.0000000000000764](https://doi.org/10.1097/MLR.0000000000000764)] [Medline: 28671930]
41. Kharrazi H, Lasser EC, Yasnoff WA, Loonsk J, Advani A, Lehmann HP, et al. A proposed national research and development agenda for population health informatics: Summary recommendations from a national expert workshop. *J Am Med Inform Assoc* 2017 Dec;24(1):2-12 [FREE Full text] [doi: [10.1093/jamia/ocv210](https://doi.org/10.1093/jamia/ocv210)] [Medline: 27018264]
42. Kharrazi H, Hatef E, Lasser E, Woods B, Rouhizadeh M, Kim J, et al. A Guide to Using Data from Johns Hopkins Electronic Health Record: Identifying Research Needs and Assessing the Availability of Behavioral/Social Variables. 2018 Sep. URL: https://ictr.johnshopkins.edu/wp-content/uploads/Phase2.Epic_Social.Guide_2018.06.30_final.pdf [accessed 2019-02-18] [WebCite Cache ID 76GvqT356]
43. Kharrazi H, Weiner JP. IT-enabled community health interventions: Challenges, opportunities, and future directions. *EGEMS (Wash DC)* 2014;2(3):1117 [FREE Full text] [doi: [10.13063/2327-9214.1117](https://doi.org/10.13063/2327-9214.1117)] [Medline: 25848627]
44. Dixon BE, Kharrazi H, Lehmann HP. Public health and epidemiology informatics: Recent research and trends in the United States. *Yearb Med Inform* 2015 Aug 13;10(1):199-206 [FREE Full text] [doi: [10.15265/IY-2015-012](https://doi.org/10.15265/IY-2015-012)] [Medline: 26293869]

45. Dixon BE, Pina J, Kharrazi H, Gharghabi F, Richards J. What's past is prologue: A scoping review of recent public health and global health informatics literature. *Online J Public Health Inform* 2015;7(2):e216 [[FREE Full text](#)] [doi: [10.5210/ojphi.v7i2.5931](https://doi.org/10.5210/ojphi.v7i2.5931)] [Medline: [26392846](#)]
46. Gamache R, Kharrazi H, Weiner JP. Public and population health informatics: The bridging of big data to benefit communities. *Yearb Med Inform* 2018 Aug;27(1):199-206 [[FREE Full text](#)] [doi: [10.1055/s-0038-1667081](https://doi.org/10.1055/s-0038-1667081)] [Medline: [30157524](#)]

Abbreviations

- A:** ICD9-Annotation
- AP:** ICD9-Annotation-Post
- B:** basic
- C:** Is-ICD9-Code
- CRF:** conditional random field
- cTAKES:** clinical Text Analysis Knowledge Extraction System
- E:** Entity-Attributes
- EHR:** electronic health record
- ICD:** International Classification of Diseases
- IRB:** Institutional Review Board
- L-BFGS:** limited-memory Broyden-Fletcher-Goldfarb-Shanno
- NLP:** natural language processing
- OOV:** out-of-vocabulary
- S:** stem
- SDH:** social determinants of health
- SGD:** stochastic gradient descent
- ShARe/CLEF:** Shared Annotated Resources/Conference and Labs of the Evaluation Forum
- T:** section
- U:** Is-Medical-Unit

Edited by C Lovis; submitted 10.12.18; peer-reviewed by S Mehrabi, M Torii, A Yates, E Kirkendall, M Laurent, Y Wang; comments to author 03.01.19; revised version received 18.02.19; accepted 07.03.19; published 26.03.19.

Please cite as:

Chen T, Dredze M, Weiner JP, Hernandez L, Kimura J, Kharrazi H

Extraction of Geriatric Syndromes From Electronic Health Record Clinical Notes: Assessment of Statistical Natural Language Processing Methods

JMIR Med Inform 2019;7(1):e13039

URL: <http://medinform.jmir.org/2019/1/e13039/>

doi: [10.2196/13039](https://doi.org/10.2196/13039)

PMID: [30862607](https://pubmed.ncbi.nlm.nih.gov/30862607/)

©Tao Chen, Mark Dredze, Jonathan P Weiner, Leilani Hernandez, Joe Kimura, Hadi Kharrazi. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Implementation of National Health Informatization in China: Survey About the Status Quo

Chen Li^{1,2*}, PhD; Xiangdong Xu^{3*}, MA; Guanghua Zhou³, MA; Kai He^{1,2}, MA; Tianliang Qi^{1,2}, BA; Wei Zhang⁴, PhD; Feng Tian^{1,2}, PhD; Qinghua Zheng^{1,2}, PhD; Jianping Hu³, MA

¹School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China

²Shanxi Province Key Laboratory of Satellite and Terrestrial Network Technology Research and Development, Xi'an Jiaotong University, Xi'an, China

³Department of Information Technology, Center of Statistics and Health Informatics, National Health Commission of People's Republic of China, Beijing, China

⁴College of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

*these authors contributed equally

Corresponding Author:

Jianping Hu, MA

Department of Information Technology

Center of Statistics and Health Informatics

National Health Commission of People's Republic of China

No. 1, West Straight Gate South Road

Beijing,

China

Phone: 86 010 68792572

Email: hujp@nhfpc.gov.cn

Abstract

Background: The National Health and Family Planning Commission (NHFPC) in China organized a nationwide survey to investigate the informatization in hospitals and regional Health and Family Planning Commissions (HFPCs) in 2017. The survey obtained valid results from 79.69% (2021/2536) of major hospitals and 81% (26/32) of provincial and 73.1% (307/420) of municipal HFPCs. The investigated topics covered hardware infrastructure, information resources, applications, systems, and organizations in health informatics.

Objective: This study aimed to provide evidence collected from the survey regarding China's health informatization and assist policy making regarding health informatics in the 13th Five-Year Plan of China.

Methods: Based on the survey, the paper presented the status quo of China's health informatization and analyzed the progress and potential problems in terms of the country's health information development policies.

Results: Related policies have helped to construct 4-level information platforms and start converging the regional data to the 3 centralized databases. The principle of informatics has been transiting from finance-centered to people-centered. Alternatively, the quality, usability, and interoperability of the data still need to be improved.

Conclusions: The nationwide survey shows that China's health informatization is rapidly developing. Current information platforms and databases technically support data exchanging between all provinces and cities. As China is continuing to improve the infrastructure, more advanced applications are being developed upon it.

(*JMIR Med Inform* 2019;7(1):e12238) doi:[10.2196/12238](https://doi.org/10.2196/12238)

KEYWORDS

health policy; health informatization; electronic health record

Introduction

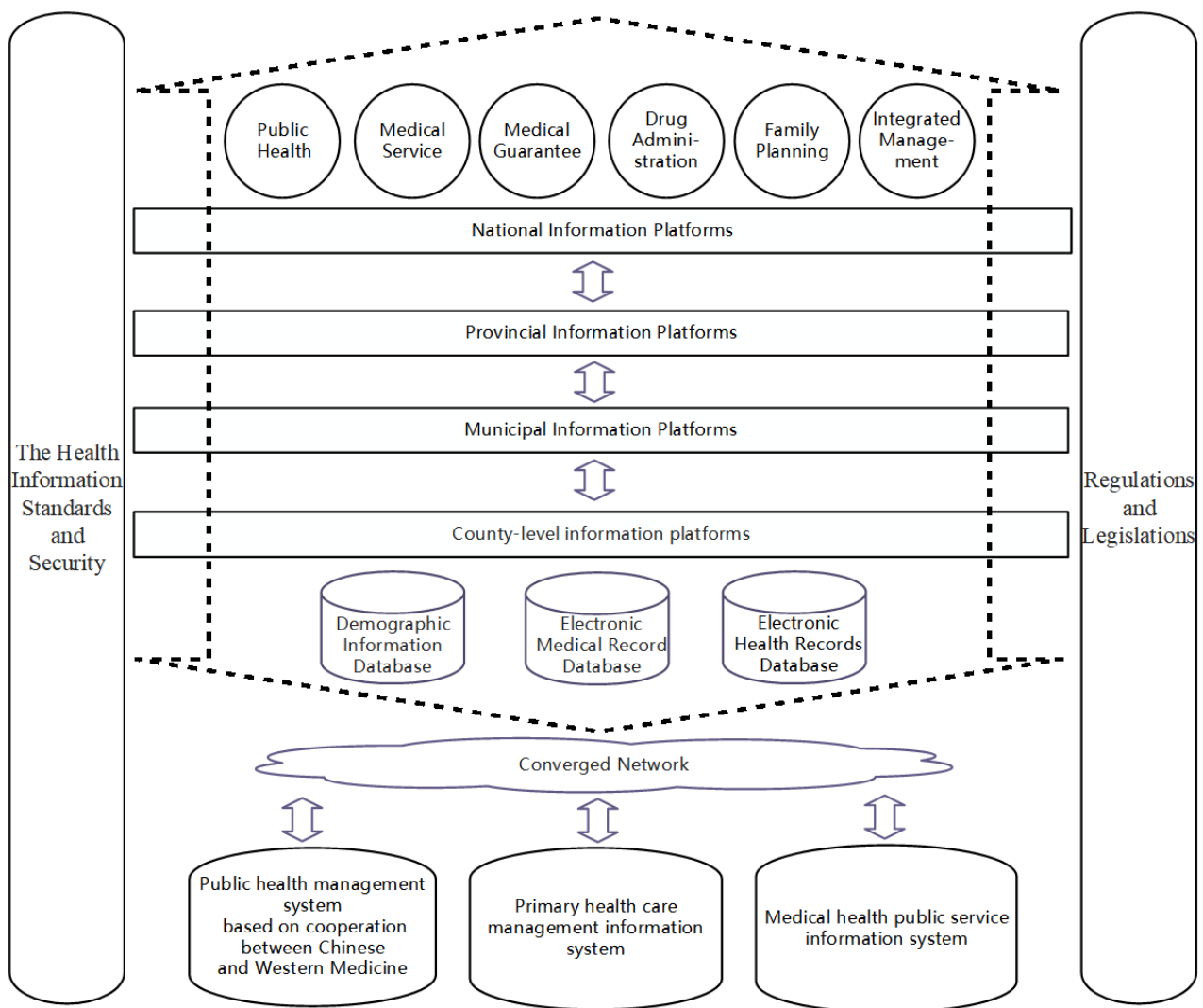
In May 2011, the 12th Five-Year Plan and its corresponding subplans made by the State Council of the People's Republic of China (PR China) were approved. One of the subplans,

Instructions for Accelerating Population Health Informatization [1], emphasized that health informatization was a crucial part of the national informatization framework and an important aspect of continuing reform of medical health in China. The difficulty of making a comprehensive plan of health

informatization for a country with 1.4 billion people was probably beyond imagination. In China, the National Health and Family Planning Commission was responsible for proposing the plan and the related policies. In 2010, the NHFPC initialized the 3521 Project, which was known as China's health informatization road map and later upgraded to the 4631-2 Project in 2013 [1,2]. Figure 1 illustrates the planned 4631-2 Project and its numeric representations:

- The 4 stands for the 4-level information platforms, which include the national information platforms, provincial information platforms, municipal information platforms, and county-level information platforms. On each platform, the medical data from different regions are integrated and shared.
- The 6 stands for 6 primary types of applications, including public health, medical service, medical guarantee, drug administration, family plan, and integrated management. The 6 types of applications are deployed on each of the 4-level information platforms.
- The 3 stands for 3 demographic health information databases (DHID) constructed on each of the 4-level information platforms, including a demographic information database (DID), electronic medical record database (EMRD), and electronic health record database (EHRD) [3]. The DID contains fundamental population information, family planning service management information, and floating population management information, etc. The EMRD stores all information from the electronic medical record. The EHRD regards the residents' personal health information as the key and stores all information from the electronic health record (EHR), which is defined as "digitally stored health care information about an individual's lifetime with the purpose of supporting continuity of care, education, and research and ensuring confidentiality at all times" [4].
- The 1 stands for one efficient and centralized network covering all kinds of health care institutions at all levels (including Chinese medical institutions).
- The 2 stands for health information standards, security, regulations, and legislation.

Figure 1. The 4631-2 Project framework of health informatization.



In October 2016, the State Council of PR China issued the Plan for Healthy China, in which the importance of informatization in health care was reemphasized [5]. Subsequently, the NHFPC released the National Population Health Informatization Development Plan (NPHIDP13) in January 2017 containing the fundamental policies [6]. At the beginning of the NPHIDP13, NHFPC conducted a nationwide survey requiring the participation of thousands of hospitals and all provincial and municipal health and family planning commissions (HFPCs). Based on the survey and its analysis, the paper discussed health informatization progress in China and its challenges. It showed that although the country's health informatization had been initialized relatively later than many countries, it has been developing at a rapid and steady pace. The initial goal of constructing the EHR systems and sharing data among regions had been achieved. With the construction of the EHRD, the development of health informatization had started transiting from finance-centered to people-centered. Some prominent problems remain: low data quality, data usability, and interoperability are yet to be solved, and the national structure of 4-level information platforms has been really challenging. The survey investigated the progress of health informatization and would support policy- and strategy-making. To our knowledge, this paper is the first on research about health informatization based on a nationwide survey in China, and similar research is rare.

Methods

The survey consists of 3 comprehensive questionnaires including the Hospital Informatization Questionnaire, the Regional HFPC Informatization Questionnaire, and the Big Data Technology in Health Informatization Questionnaire. Each questionnaire was developed by a group of 8 experts and covers a range of key aspects of informatization, including administration, applications, hardware infrastructure, big data applications, information exchange platforms, information security, and related standards. In all, 32 provincial HFPCs, 420 municipal HFPCs, and 2536 major hospitals were invited. Valid responses came from 81% (26/32) of provincial HFPCs, 73.1% (307/420) of municipal HFPCs, and 79.69% (2021/2536) of hospitals.

The NHFPC is the state administrative department for health-related affairs in the State Council of PR China. The provincial and municipal HFPCs are local district branches under the supervision of the NHFPC. A provincial HFPC was responsible for the health affairs of the corresponding province, including ensuring public health; implementing health policies; promoting health care reform; and managing affiliated medical administrations, hospitals, and schools. Similarly, the municipal HFPC has the same responsibility at the municipal level. Among 2021 hospital respondents, there were 42 NHFPC-affiliated hospitals, 371 provincial NHFPC-affiliated hospitals, and 1608 municipal NHFPC-affiliated hospitals.

Similar research regarding health informatization has been carried out in a number of countries [7-10]. Compared with this research, our information was collected from an official survey conducted by the country's health ministry (NHFPC), and the survey covered a wider range about health informatization,

including hardware and software of health infrastructure, national data exchanging platforms, health informatization applications, investment, and staffing.

Results

Three Phases of Informatization

The 4631-2 Project, the guidance of health informatization in China, consisted of 3 phases. The first phase was to improve health information technology (IT) hardware infrastructure and the application systems of routine businesses and set up the information standards and security systems. The second phase was to construct the information platform levels according to region levels. The hierarchy and mechanism enabled data exchange between different regions and applications. The third phase was to apply big data technologies, artificial intelligence (AI), and Internet Plus (China's plan to manage its information superhighway) proposed in a government work report from March 5, 2015, that refers to the application of the internet and other IT in conventional industries [11] and enables advanced applications (eg, predictive modeling, clinical decision support, disease or safety surveillance, public health, and research) [2,12].

People-Centered Health Information Infrastructure

The various medical information systems, such as EHR systems, were the fundamental data resources. The continuously improved EHR systems and 3 DHIDs converged the data and gradually became the central resources. The NHFPC had promoted nationwide deployment of EHR systems and regular data synchronization into the EHRD to compile a complete EHR for each citizen regardless of where the original data were generated. Each region's local HFPC was responsible for the implementation of the EHRD. The survey showed that 86% (18/22) of provincial HFPCs and 75.6% (232/307) of municipal HFPCs had established the EHRD as of March 2017. Meanwhile, 71% (15/21) of provincial HFPCs and 70.4% (216/307) of municipal HFPCs had EHRs for 70% population in the corresponding regions. In addition, 16.19% (3,880,879/23,956,343) of provincial and 14.31% (271,920/1,900,209) of municipal HFPC EHRs were built in 2016 alone.

Data from 6 types of applications and related systems were organized into the other 2 core databases, the DID and EMRD. The survey showed that 96% (22/23) of provincial HFPCs had built DIDs and 48% (10/21) of provincial and 31.6% (97/307) of municipal HFPCs had started constructing EMRDs. The NPHIDP13 had explicitly required the data from 6 applications and other related systems to be available to 3 DHIDs. It had also required that the 3 DHIDs should be integrated following the people-centered principle. Meanwhile, NPHIDP13 had explicit criteria about system consistency, accuracy, and integrity.

Data Exchanging on the Four-Level Information Platforms

Centralized medical data had significant potential for applications and research. Enabling information sharing within and across regions had been attempted by a few countries and

was demonstrated to be difficult for various reasons [13-17]. The 4631-2 Project supported regional informatization that collected data from 6 primary types of applications, 3 population health information databases (PHIDs), and other public health systems. Systematic informatization also promoted sharing data within and across regions. As the development of 4-level information platforms was the core task of regional informatization, the platforms' construction had been one of the key tasks in the Instructions for Accelerating Population Health Informatization [1]. In the 4-level information platforms, most data were generated at the provincial and municipal platforms, so these 2 levels of the platforms naturally became the focuses of the survey. As of March 2017, 32% (7/22) of provincial and 35.2% (108/307) of municipal HFPCs had built the information platforms. All provincial and all but 17 (6%) of the municipal HFPCs had started establishing information platforms. More importantly, all 32 provincial platforms had been connected with the national platforms.

The 4-level information platforms offered 2 advantages. The first was centralization of storing and sharing health information. The 4-level information platforms also connected public health information systems with medical and health service institutions within regions. The integration and sharing within the information platforms would directly improve the management efficiency of health administrations, for example, by providing panoramic information of the entire country during an epidemic breakout. The second was to unify management for EHRs. Based on the information platforms, unified management could facilitate various nationwide standardization.

Advanced Applications

Advanced applications based on the converged data and technologies like big data, AI, and Internet Plus were considered to have great potential to benefit the people and the health care industry [18-20]. The NPHIDP13 plan had been designed with this consideration. The survey showed that 67% (16/24) of provincial and 31.8% (100/314) of municipal HFPCs and 32.35% (647/2000) of hospitals had implemented various big data applications. Business intelligence reports, medical quality control warnings, clinical medical records, and related data retrieval had been well received by the users. A poll within the survey showed that clinical quality control and chronic disease management based on big data were most appreciated by medical professionals. At the same time, the poll showed that 1180 (59.00%), 1062 (53.10%), and 944 (47.20%) of the 2000 hospitals expected more funds, consistent standards, and interinstitution cooperation, respectively, regarding big data research and application development.

The combination of big data, AI, and Internet Plus still needed to overcome technical and technological challenges to reveal the true potential and support the most demanded applications (eg, precision medicine and AI-assisted clinical decision support systems) [21-23]. One of the issues was the collection and processing of multimodel data from various sources. The NPHIDP13 had tried to structure the data as much as possible. The Big Data Technology in Health Informatization Questionnaire showed that 45.10% (902/2000) of hospitals had

structured health big data, 41.40% (828/2000) had semistructured electronic medical record health big data, 28.50% (570/2000) had unstructured medical image data, and 6.00% (120/2000) had unstructured omics data. For regional HFPCs, 24% (5/21) of provincial and 15.9% (50/314) of municipal HFPCs had unstructured medical image data, and 48% (10/21) of provincial and 1.9% (6/314) of municipal HFPCs had unstructured omics data. Meanwhile, the unstructured data, which could be 80% or 85% of overall data, needed more professional curating staff and technologies like natural language processing and image processing to automatically process it [24,25].

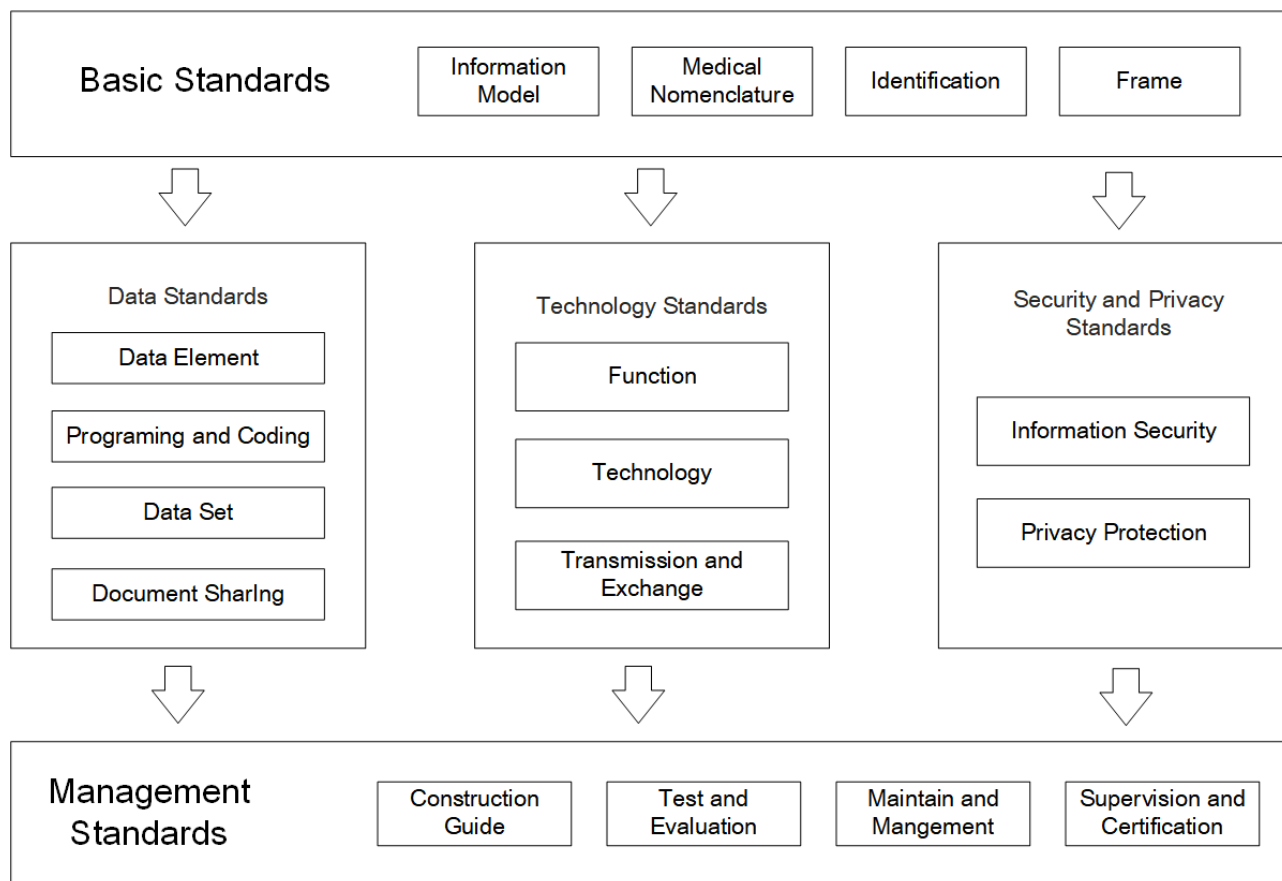
Investment and Staffing

The national health informatization project was extremely complex and expensive, but the Chinese government played an essential and supportive role in the project: increases were not seen in local medical service prices and data privacy and usability were better assured. As of March 2017, the 32 provincial platforms received ¥1.87 billion (US \$276.6 million) from the central government (not including funds from local governments). During the 12th Five-Year Plan, the central government invested more than ¥10 billion (US \$1.5 billion) into health informatization [1,3,6].

Since health care reform in 2009, the number of doctors (approximately 1,730,000 in 2015) has increased by 3.3% annually. The annual rate of new nurses is even larger (9.9%, approximately 472,000 in 2015). The increasing rate of pharmacists is 3.1% (approximately 109,000 in 2015) [26-28]. In comparison with the increasing number of medical professionals, the number of IT professionals is relatively low. Often, the responsibility of the IT departments must be limited in coordinating the needs with the health IT service providers due to a limited number of IT staff. According to the survey, 19% (4/21) of provincial and 85.0% (267/314) of municipal HFPCs had fewer than 9 IT staff members. For the hospitals, only 7.00% (140/2000) had more than 20 full-time IT staff members; the number of IT staff was not enough to meet the actual work needs. The survey showed that 76.00% (1520/2000) of medical institutions expressed the lack of IT staff and the difficulty of recruitment.

Standards and Certification

Standards, motivation, and credibility were 3 keys to effective interoperability of health information technology [29]. Standards provided the possibility of efficient communications, motivation promoted the usability and consequently created more data, and credibility ensured a trustworthy environment for using data. China's information standardization framework comprises 3 parts: data standards, technology standards, and security and privacy standards (Figure 2). Since 2010, 283 national health informatization standardization projects had been approved and initiated, ranging from data acquisition and exchange to information management, storage, cataloging, and security, etc. Among the approved standards, 33 have been adopted by consortiums. Of 250 planned industry standards, 209 have been published, 16 will be published soon, another 16 have passed preliminary assessment, and 10 are under development.

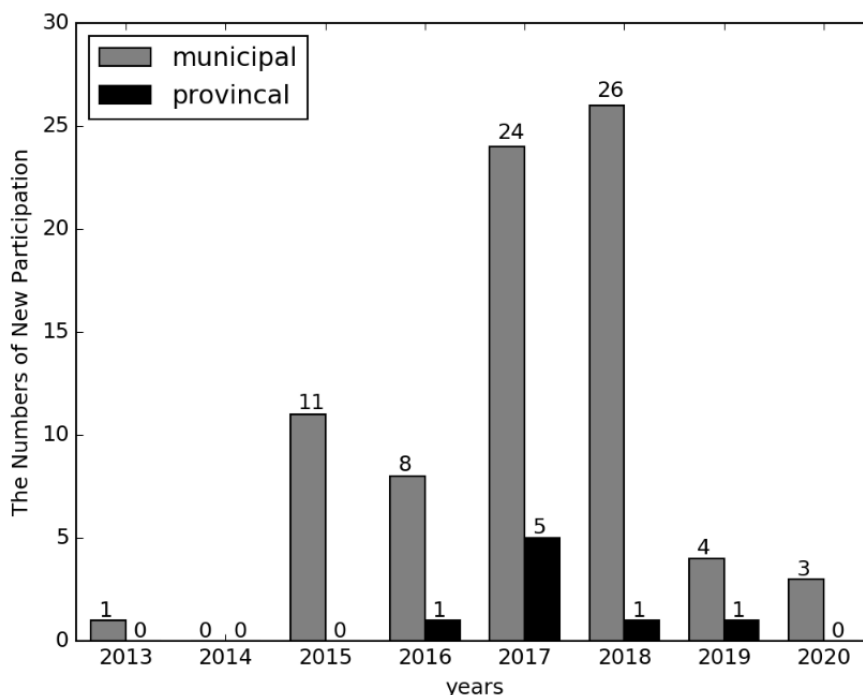
Figure 2. Framework for information standardization.

Informatization standards not only cover hardware-related issues such as data acquisition, storage, and transmission, called hard standards, but also software-related issues such as metadata standards, information usage standards, supervision and certification, called soft standards. Generally, soft standards are relatively immature compared with hard standards. Furthermore, hard and soft standards should be better integrated within the information standardization framework (Figure 2). Many existing information standards still play separate roles, forming the biggest barriers for interoperability. With awareness of these issues, the NHFPC is making efforts to improve the synergy of health informatization standards to enhance interoperability.

To help the regional HFPCs and hospitals stay informed about standardization and better implementation of standards, a series

of certification programs has been launched by the NHFPC. The National Medical Health Information System's Interconnection and Interoperability Standardization Certification was started in 2013. The certification covers a wide range of data standards, document sharing, platform functions, hardware infrastructure, information security, business applications, etc [30]. As shown by the survey, 9.1% (24/263) of municipal platforms participated in the certification in 2017. Because the program for provincial platforms just started, only one provincial platform has participated. Meanwhile, 46% (8/17) of provincial and 31.3% (77/246) of municipal HFPCs have a clear timeline for taking part in the program. Figure 3 shows the regional HFPCs' planned schedule of participating the program.

Figure 3. Schedule of regional Health and Family Planning Commissions obtaining certifications.



Discussion

Principal Findings

Compared with 59.9% of eligible nonprofit hospitals that get funding in the United States [31], almost all major hospitals in China are public and nonprofit. The Chinese government is responsible for funding support and therefore plays a more active role in health informatization.

One of the key achievements in recent years was the development of nationwide 4-level information platforms. It was an ambitious and challenging task that established a certain level of connectivity between HFPCs in the country. This might enhance information sharing and collaboration and may directly support allocating health care resources according to regional differences. For medical professionals, it largely improved information accessibility—for example, to be able to make rapid diagnoses based on more complete information or reduce the number of reexamination cases. For patients, it improved the accessibility of health information and enabled interregional medical health and insurance services. For health administrations, it delivered real-time intelligence of dynamic situations about health service resources to support scientific management, such as inspecting and maintaining reasonable

medical expenses. The potential of converged data was still far from being fully exploited and many more data applications were still under development. The survey showed that 71% of provincial HFPCs regarded developing and improving applications as a main task for the future.

Another more serious problem is low data quality. According to the analysis of the survey, empty or invalid data, disorders of disease description language, lack of retrieval tools, lack of metadata management, and inability to handle the text data of electronic medical records are prominent problems in China’s health informatization. Only 66.8% of municipal NFPCs implements or plans to implement data quality management system standards. The problem is serious and may cause unexpected barriers for the usability of the 3 databases and the platform. Without high-quality local data, centralized medical data on the 4-level information platforms won’t play a real role and advanced applications based on converged data and technologies like big data and AI will not have a solid foundation. The survey also reported that 74% of provincial HFPCs put this at the top of the list of problems to be solved. To summarize, we list the best practices and corresponding barriers and successes of the national health informatization implementation in Table 1.

Table 1. Best practices and corresponding main barriers and successes in the implementation of national health informatization.

Best practices	Main barriers	Successes
Three demographic health information databases	Low data quality	Large enough coverage for population
Four-level information platforms	Difficultly exchanging data	National health data exchanging
Advanced informationalized applications	Unstructured data processing	Big data, AI, and Internet Plus applications
Established standards	Lack of synergy of standards	A total of 283 national standardizations were approved

Conclusions

Health informatization in China started in the relative recent years and has developed at a very rapid pace. Its objective is to enable advanced clinical research and applications at the beginning while the IT infrastructure is built. Initially, informatization aimed to improve efficiency and administration of finances. Gradually, informatization transitioned from finance-centered to people-centered in order to bring full potential in terms of health outcomes. The transition was naturally occurring while the goal of informatization was explicitly set to bring more benefits for every patient in NPHIDP13.

According to NPHIDP13, personalized health care services and products based on perception technology will start being applied and promoted by 2020; DHIDs storing individual lifelong records for the entire population will be eventually built; and novel technologies, which might bring revolutionary outcomes in health care and research, were considered while designing the current systems and planning future advanced applications. Generally, NPHIDP13 planned the projects for the period of 2016 to 2020 to solve remaining problems. They include upgrading current systems with new technologies, continuing development of applications of key businesses, improving coverage of the population in 3 DHIDs, and improving interconnection and interoperability between the 4-level platforms.

Acknowledgments

This work has been supported by the National Key Research and Development Program of China (2018YFC0910404); National Natural Science Foundation of China (61772409); Ministry of Education–Research Foundation of China Mobile Communication Corp (MCM20160404); the consulting research project of the Chinese Academy of Engineering (The Online and Offline Mixed Educational Service System for “The Belt and Road” Training in MOOC China); Project of China Knowledge Centre for Engineering Science and Technology; the innovation team from the Ministry of Education (IRT_17R86); and the Innovative Research Group of the National Natural Science Foundation of China (61721002).

Conflicts of Interest

None declared.

References

1. National Health and Family Planning Commission of the People's Republic of China. Instructions for Accelerating Population Health Informatization. URL: <http://www.nhfpc.gov.cn/guihuaxxs/s10741/201312/09bce5f480e84747aa130428ca7fc8ad.shtml> [accessed 2018-09-17] [WebCite Cache ID 72VEKqqaj]
2. China Digital Medical Network. The 4631-2 framework of the top level design of national health family resources. URL: <http://www.hc3i.cn/zhuanti/2014litf/> [accessed 2018-09-17] [WebCite Cache ID 72VERKeza]
3. National Health and Family Planning Commission of the People's Republic of China. Questions and answers about Instructions for Accelerating Population Health Informatization. URL: <http://www.nhfpc.gov.cn/guihuaxxs/s10742/201312/2519dea9a4b14318a0736881116275ee.shtml> [accessed 2018-09-17] [WebCite Cache ID 72VEXL2OG]
4. Iakovidis I. Towards personal health record: current situation, obstacles and trends in implementation of electronic healthcare record in Europe. *Int J Med Inform* 1998;52(1-3):105-115. [Medline: 9848407]
5. CPC Central Committee, State Council. The plan for Healthy China 2030. 2016. URL: http://www.gov.cn/xinwen/2016-10/25/content_5124174.htm [accessed 2018-09-17] [WebCite Cache ID 72VEuZz0V]
6. National Health and Family Planning Commission of the People's Republic of China. National Population Health informatization development plan for 13th Five-Year Plan. URL: <http://www.moh.gov.cn/guihuaxxs/s10741/201702/ef9ba6fbe2ef46a49c333de32275074f.shtml> [accessed 2018-10-27] [WebCite Cache ID 73TKac8r6]
7. Ludwick DA, Doucette J. Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. *Int J Med Inform* 2009 Jan;78(1):22-31. [doi: 10.1016/j.ijmedinf.2008.06.005] [Medline: 18644745]
8. Deutsch E, Duftschmid G, Dorda W. Critical areas of national electronic health record programs—is our focus correct? *Int J Med Inform* 2010 Mar;79(3):211-222. [doi: 10.1016/j.ijmedinf.2009.12.002] [Medline: 20079685]
9. Jha AK, Doolan D, Grandt D, Scott T, Bates DW. The use of health information technology in seven nations. *Int J Med Inform* 2008 Dec;77(12):848-854. [doi: 10.1016/j.ijmedinf.2008.06.007] [Medline: 18657471]
10. Chen C, Hsu C, Worley J. Evaluation systems of the electronic medical records: a comparative study. 2013 Presented at: 2013 4th IEEE International Conference on Software Engineering and Service Science; 2013; Beijing.
11. State Council of the People's Republic of China. China unveils Internet Plus action plan to fuel growth. URL: http://english.gov.cn/policies/latest_releases/2015/07/04/content_281475140165588.htm [accessed 2019-01-10] [WebCite Cache ID 72VHFG9qS]
12. Report on the work of the government at the third meeting of the Twelfth National People's Congress. URL: <http://sh.people.com.cn/n2/2017/0317/c138654-29868849.html> [accessed 2019-01-10] [WebCite Cache ID 72VrCWivW]

13. Blumenthal D. Wiring the health system—origins and provisions of a new federal program. *N Engl J Med* 2011 Dec 15;365(24):2323-2329. [doi: [10.1056/NEJMsr1110507](https://doi.org/10.1056/NEJMsr1110507)] [Medline: [22168647](https://pubmed.ncbi.nlm.nih.gov/22168647/)]
14. Thomson S, Osborn R, Squires D. International profiles of health care systems. 2013. URL: https://www.commonwealthfund.org/sites/default/files/documents/media_files_publications_fund_report_2013_nov_1717_thomson_intl_profiles_hlt_care_sys_2013_v2.pdf [accessed 2019-01-10] [WebCite Cache ID 72VHbRc7I]
15. Fragidis L, Chatzoglou P. Development of nationwide electronic health record (NEHR): an international survey. *Health Policy Technol* 2017 Jun;6(2):124-133 [FREE Full text] [doi: [10.1016/j.hlpt.2017.04.004](https://doi.org/10.1016/j.hlpt.2017.04.004)]
16. Hofmarcher MM, Quentin W. Austria: health system review. *Health Syst Transit* 2013;15(7):1 [FREE Full text] [Medline: [24334772](https://pubmed.ncbi.nlm.nih.gov/24334772/)]
17. Stroetmann K, Artmann J, Stroetmann V, Protti D, Dumortier J, Giest S, et al. Office for Official Publications of the European Communities. 2011. European countries on their journey towards national eHealth infrastructures URL: http://www.ehealthnews.eu/images/stories/pdf/ehstrategies_final_report.pdf [accessed 2018-09-18] [WebCite Cache ID 72VscbIxF]
18. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* 2013 Apr 3;309(13):1351-1352. [doi: [10.1001/jama.2013.393](https://doi.org/10.1001/jama.2013.393)] [Medline: [23549579](https://pubmed.ncbi.nlm.nih.gov/23549579/)]
19. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3 [FREE Full text] [doi: [10.1186/2047-2501-2-3](https://doi.org/10.1186/2047-2501-2-3)] [Medline: [25825667](https://pubmed.ncbi.nlm.nih.gov/25825667/)]
20. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
21. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff (Millwood)* 2014 Jul;33(7):1123-1131. [doi: [10.1377/hlthaff.2014.0041](https://doi.org/10.1377/hlthaff.2014.0041)] [Medline: [25006137](https://pubmed.ncbi.nlm.nih.gov/25006137/)]
22. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015 Feb 26;372(9):793-795. [doi: [10.1056/NEJMp1500523](https://doi.org/10.1056/NEJMp1500523)] [Medline: [25635347](https://pubmed.ncbi.nlm.nih.gov/25635347/)]
23. Frost S. Drowning in big data? Reducing information technology complexities and costs for healthcare organizations. URL: <http://www.emc.com/collateral/analyst-reports/frost-sullivan-reducing-information-technology-complexities-ar.pdf> [accessed 2018-09-18] [WebCite Cache ID 72VtxHIUE]
24. Grimes S. Clarabridge BridgePoints. 2008. Unstructured data and the 80 percent rule URL: <http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/> [accessed 2018-09-18] [WebCite Cache ID 72VuoU2Dy]
25. Li C, Liakata M, Rebholz-Schuhmann D. Biological network extraction from scientific literature: state of the art and challenges. *Brief Bioinform* 2014 Sep;15(5):856-877. [doi: [10.1093/bib/bbt006](https://doi.org/10.1093/bib/bbt006)] [Medline: [23434632](https://pubmed.ncbi.nlm.nih.gov/23434632/)]
26. López FJC, Arenillas LJ, Calleja PS, Botia PE, Casado NI, Deyá AE, et al. Health care resources for stroke patients in Spain, 2010: analysis of a national survey by the Cerebrovascular Diseases Study Group. *Neurologia* 2011 Oct;26(8):449-454 [FREE Full text] [doi: [10.1016/j.nrl.2010.10.014](https://doi.org/10.1016/j.nrl.2010.10.014)] [Medline: [21440962](https://pubmed.ncbi.nlm.nih.gov/21440962/)]
27. National Health and Family Planning Commission of the People's Republic of China. *China Health and Family Planning Statistical Yearbook*. Beijing: Peking Union Medical and College Publishing House; 2016.
28. Ministry of Health of the People's Republic of China. *China Health Statistical Yearbook*. Beijing: Peking Union Medical and College Publishing House; 2009.
29. Office of the National Coordinator for Health IT. Connecting health and care for the nation? A shared nationwide interoperability roadmap. URL: <https://www.healthit.gov/sites/default/files/hie-interoperability-nationwide-interoperability-roadmap-final-version-1.0.pdf> [accessed 2019-01-10] [WebCite Cache ID 72VvqTyDI]
30. National Health and Family Planning Commission of the People's Republic of China. National medical health information interoperability standardization maturity certification. 2017. URL: <http://www.nhfpc.gov.cn/mohwsbwstjxxzx/s8553/201709/06ac0b5228674ce487405de8e939d40b.shtml> [accessed 2018-09-18] [WebCite Cache ID 72VvuSCv5]
31. Adler-Milstein J, Jha AK. HITECH Act drove large gains in hospital electronic health record adoption. *Health Aff (Millwood)* 2017 Aug 01;36(8):1416-1422. [doi: [10.1377/hlthaff.2016.1651](https://doi.org/10.1377/hlthaff.2016.1651)] [Medline: [28784734](https://pubmed.ncbi.nlm.nih.gov/28784734/)]

Abbreviations

- AI:** artificial intelligence
- DHID:** demographic health information database
- DID:** demographic information database
- EHR:** electronic health record
- HFPC:** Health and Family Planning Commission
- EHRD:** electronic health record database
- EMRD:** electronic medical record database
- IT:** information technology
- NHFPC:** National Health and Family Planning Commission

NPHIDP13: National Population Health Informatization Development Plan

PHID: population health information database

PR China: People's Republic of China

Edited by G Eysenbach; submitted 18.09.18; peer-reviewed by L Fragidis, E Andrikopoulou; comments to author 24.10.18; revised version received 06.11.18; accepted 09.12.18; published 21.02.19.

Please cite as:

Li C, Xu X, Zhou G, He K, Qi T, Zhang W, Tian F, Zheng Q, Hu J

Implementation of National Health Informatization in China: Survey About the Status Quo

JMIR Med Inform 2019;7(1):e12238

URL: <http://medinform.jmir.org/2019/1/e12238/>

doi: [10.2196/12238](https://doi.org/10.2196/12238)

PMID: [30789350](https://pubmed.ncbi.nlm.nih.gov/30789350/)

©Chen Li, Xiangdong Xu, Guanghua Zhou, Kai He, Tianliang Qi, Wei Zhang, Feng Tian, Qinghua Zheng, Jianping Hu. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 21.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

An Assessment of the Interoperability of Electronic Health Record Exchanges Among Hospitals and Clinics in Taiwan

Hsyien-Chia Wen¹, PhD; Wei-Pin Chang¹, PhD; Min-Huei Hsu², MD, PhD; Cheng-Hsun Ho³, PhD; Chi-Ming Chu⁴, PhD

¹School of Healthcare Administration, College of Management, Taipei Medical University, Taipei, Taiwan

²Graduate Institute of Bio-medical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan

³Graduate Institute of Information Management, National Taipei University, Taipei, Taiwan

⁴Section of Biostatistics and Bioinformatics, Department of Epidemiology, School of Public Health, National Defense Medical Center, Taipei, Taiwan

Corresponding Author:

Hsyien-Chia Wen, PhD

School of Healthcare Administration

College of Management

Taipei Medical University

10F., No 172-1

Section 2, Keelung Road, Xinyi District

Taipei, 106

Taiwan

Phone: 886 6638 2736 ext 1019

Email: hcwenn@tmu.edu.tw

Abstract

Background: The rapid aging of the Taiwanese population in recent years has led to high medical needs for the elderly and increasing medical costs. Integrating patient information through electronic health records (EHRs) to reduce unnecessary medications and tests and enhance the quality of care has currently become an important issue. Although electronic data interchanges among hospitals and clinics have been implemented for many years in Taiwan, the interoperability of EHRs has not adequately been assessed.

Objective: The study aimed to analyze the efficiency of data exchanges and provide suggestions for future improvements.

Methods: We obtained 30 months of uploaded and downloaded data of EHRs among hospitals and clinics. The research objects of this study comprised 19 medical centers, 57 regional hospitals, 95 district hospitals, and 5520 clinics. We examined 4 exchange EHR forms: laboratory test reports, medical images, discharge summaries, and outpatient medical records. We used MySQL (Oracle Corporation) software (to save our data) and phpMyAdmin, which is a Personal Home Page program, to manage the database and then analyzed the data using SPSS 19.0 statistical software.

Results: The quarterly mean uploaded volume of EHRs among hospitals was 52,790,721 (SD 580,643). The quarterly mean downloaded volume of EHRs among hospitals and clinics was 650,323 (SD 215,099). The ratio of uploaded to downloaded EHRs was about 81:1. The total volume of EHRs was mainly downloaded by medical centers and clinics, which accounted for 53.82% (mean 318,717.80) and 45.41% (mean 269,082.10), respectively, and the statistical test was significant among different hospital accreditation levels ($F_2=7.63$; $P<.001$). A comparison of EHR download volumes among the 6 National Health Insurance (NHI) branches showed that the central NHI branch downloaded 11,366,431 records (21.53%), which was the highest, and the eastern branch downloaded 1,615,391 records (3.06%), which was the lowest. The statistical test among the 6 NHI branches was significant ($F_5=8.82$; $P<.001$). The download volumes of laboratory tests reports and outpatient medical records were 26,980,425 (50.3%) and 21,747,588 records (40.9%), respectively, and were much higher than medical images and discharge summaries. The statistical test was also significant ($F=17.72$; $P<.001$). Finally, the download time showed that the average for x-rays was 32.05 seconds, which was the longest, and was 9.92 seconds for electrocardiogram, which was the shortest, but there was no statistically significant difference among download times for various medical images.

Conclusions: After years of operation, the Electronic Medical Record Exchange Center has achieved the initial goal of EHR interoperability, and data exchanges are running quite stably in Taiwan. However, the meaningful use of EHRs among hospitals

and clinics still needs further encouragement and promotion. We suggest that the government's leading role and collective collaboration with health care organizations are important for providing effective health information exchanges.

(*JMIR Med Inform* 2019;7(1):e12630) doi:[10.2196/12630](https://doi.org/10.2196/12630)

KEYWORDS

electronic health records; interoperability; data exchange; hospitals; clinics

Introduction

Background

Many US studies have shown that electronic health records (EHRs) reduce repeated exams and medications, which can cause unnecessary costs, thereby improving patient safety and quality of care [1-8]. *The American Recovery and Reinvestment Act* of 2009 authorizes the Centers for Medicare and Medicaid Services (CMS) to award incentive payments to eligible professionals (EPs) and hospitals that demonstrate meaningful use of certified EHRs. The CMS EHR Incentive Program encourages EPs and hospitals to adopt EHRs and apply their meaningful use in 3 stages over 5 years. Clinically, the definition of meaningful use is using certified EHR technology to improve the quality, safety, and efficiency, which results in better clinical outcomes. Meaningful use sets specific objectives that EPs and hospitals must achieve to qualify for the CMS incentive programs. These specific objectives are divided into core and menu sets to check and pay those who meet these requirements, as detailed in the Clinical Quality Measure, such as electronically capturing health information in a standardized format, engaging in rigorous health information exchanges, and improving health outcomes [9].

Taiwan's EHRs were initially developed by the National Health Informatics Project in 2004. The EHR exchange was implemented in 3 stages: stage 1 (2008-2011) began and promoted the EHR plan, stage 2 (2010-2012) accelerated EHR adoption in hospitals and clinics, and stage 3 (2013-2015) subsidized the interoperability and application of EHRs. According to statistics of the Electronic Medical Record (EMR) Exchange Center (EEC) in 2016, 411 of 496 hospitals (80.4%) and about 5244 of 9782 private clinics (53.6%) were certified as having interoperable EHRs [10].

On the basis of the configuration of the EHR data exchange system in Taiwanese hospitals (Figure 1), patients are allowed to use a National Health Insurance (NHI) integrated circuit (IC) card and can ask their physician to retrieve their medical information from hospitals they have previously visited. Patients have to sign a written consent form to authorize the physician before retrieving their medical records. The EEC functions only as an EHR index generation and search service platform for hospitals and clinics. Hospitals' information systems are connected to the EEC through an EMR gateway. The hospital

converts laboratory test reports, medical images, discharge summaries, and outpatient medical records in its EMR system into standardized files and saves them on the EMR gateway. The EEC generates an index of all XML files on EMR gateways of all hospitals and provides search and retrieval services for hospitals and clinics. In this system, Health Level 7 (HL7)/Clinical Document Architecture, Release 2 (CDA R2) standards are used to generate clinical documents and the Integrating the Healthcare Enterprise (IHE) Cross-enterprise Document Sharing (XDS) profile for the communication infrastructure. The EMR gateway receives clinical documents from the hospital information system, registers the metadata in the document registry (EEC), and stores them for 6 months. An example of a patient who asked a hospital to retrieve and download his discharge summary for receiving continuous care is illustrated in Figure 2.

Some large hospitals in Taiwan have met level-6 requirements of the EMR Adoption Model, which was built by the Healthcare Information and Management Systems Society [11,12]. Taking EMR information of Taichung Veterans General Hospital (TVGH) [12] as an example, almost 90% of medical records for outpatients, emergency visits, examinations, operations, and inpatients in this medical center were digitalized. As 80%-90% of patients registered before their visit to TVGH, drug duplications and interactions can automatically be detected by the exchange records through the EEC. Therefore, it can effectively alert physicians and prevent them from prescribing unnecessary medications and treatments.

In addition to continuously implementing paperless and interoperable medical records, Taiwan launched the Smart Healthcare plan in 2009, which uses information and communications technology (ICT) to provide a Pharma Cloud System [13] to reduce therapeutic duplications and enhance public medication safety. The Personal Health Bank [14,15], a set of personal health records applied in 2014, allows patients to access their health data for various health-enhancing applications. Under these innovative services, the delivery of integrated health care in Taiwan can become more convenient and effective.

Objective

The study aimed to analyze the efficiency of data exchanges and provide suggestions for future improvements.

Figure 1. Configuration of the electronic health record data exchange system in Taiwanese hospitals. CDA: Clinical Document Architecture; DICOM: Digital Imaging and Communications in Medicine; EHR: electronic health record; HIS: hospital information system; MOHW: Ministry of Health and Welfare; NHI-VPN: National Health Insurance-Virtual Private Network.

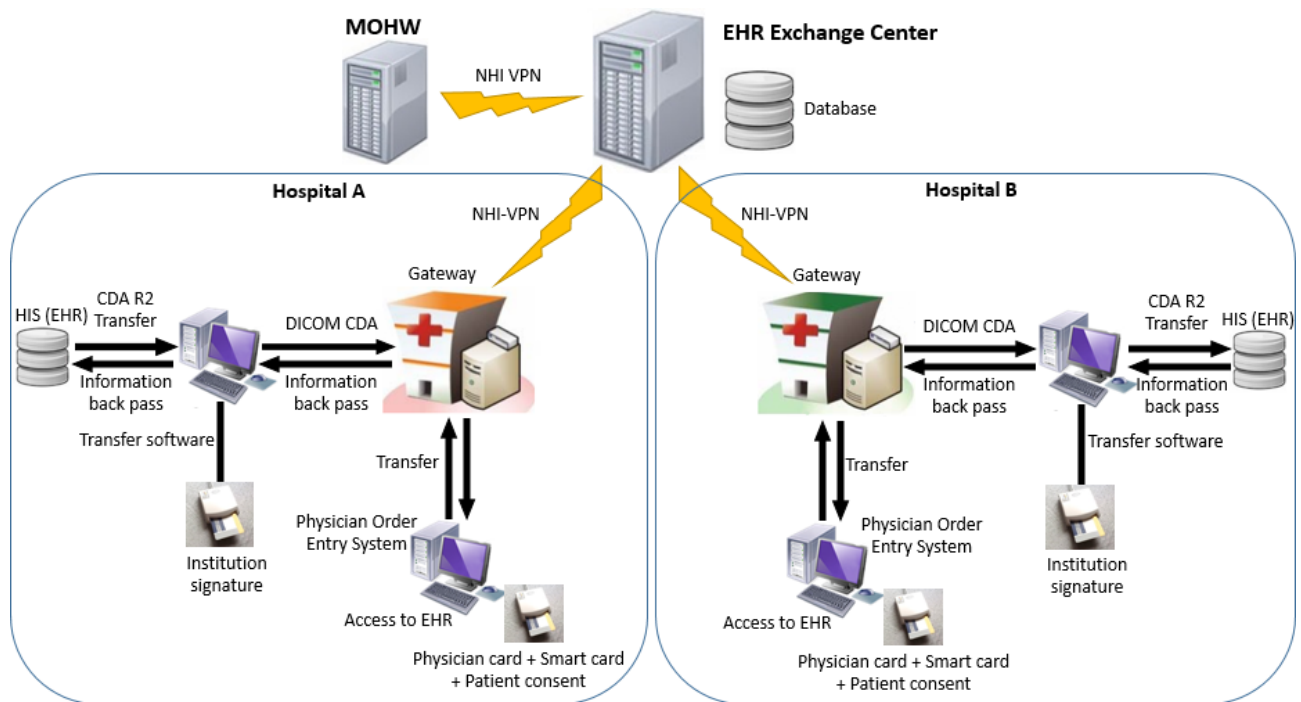


Figure 2. A retrieved discharge summary downloaded from the electronic health record exchange system.



Methods

Data of this study provided by the EEC of the Ministry of Health and Welfare consisted of 19 medical centers (>500 beds), 57 regional hospitals (>300 beds), 95 district hospitals (>100 beds), and 5520 clinics. The study period was from January 2015 to June 2017, for a total of 30 months. A total of 4 EHR exchange forms consisted of discharge summaries, outpatient records (including medication sheets), laboratory tests, and medical images. We used MYSQL software to save our data and used phpMyAdmin to manage them. Finally, we used SPSS 19.0

statistical software to analyze the variables, such as EHR exchange forms, accreditation level, geographic area, and time required for the EHRs to be downloaded.

Results

Descriptive Data of Uploaded and Downloaded Electronic Health Record Volumes

We used 3 months (1 quarter) as the analytical unit of EHR upload and download volumes, and the study period was from January 2015 to June 2017. The mean hospital uploaded EHR

volume was 52,790,721, with an SD of 580,643, a maximum of 58,546,658, and a minimum of 45,549,300 records (Figure 3). The line in Figure 3 illustrates that the volume of uploads by hospitals tended to be stable. However, the volume of downloads by hospitals fluctuated (Figure 4). The mean hospital download volume was 650,323, with an SD of 215,099, a maximum of 1,085,665, and a minimum of 59,087 records. In the first quarter of each year, the download volume was significantly lower than the average because of the lunar new year vacation. The ratio of uploaded to downloaded EHRs was about 81:1.

Variables Affecting the Uploaded Volume of Electronic Health Records

Hospital upload and download volumes of EHRs were analyzed by hospital accreditation level, NHI branch, and EHR format,

as shown in Table 1. First, average volumes of uploaded EHRs ranked by hospital accreditation of medical centers (>500 beds) with 200,927,617, regional hospitals (>300 beds) with 223,199,389, and district hospital (>100 beds) with 10,378,021 records, which reached statistical significance ($F=154.81$; $P<.001$). A comparison of the 6 NHI branches representing 6 different geographical regions of Taiwan showed that the Taipei branch had the highest average volume with 18,031,142, whereas the eastern branch had the lowest with 1,615,391 records, and the statistical test was significant ($F=360.52$; $P<.001$). Moreover, the 4 EHR exchange forms indicated that the highest volume was for laboratory tests with 26,980,425.5 records. Discharge summaries were the lowest with 675,093.5 records, and this was significant according to the statistical test ($F=887.23$; $P<.001$).

Figure 3. Electronic health record (EHR) volumes uploaded by hospitals from 2015 to 2017.

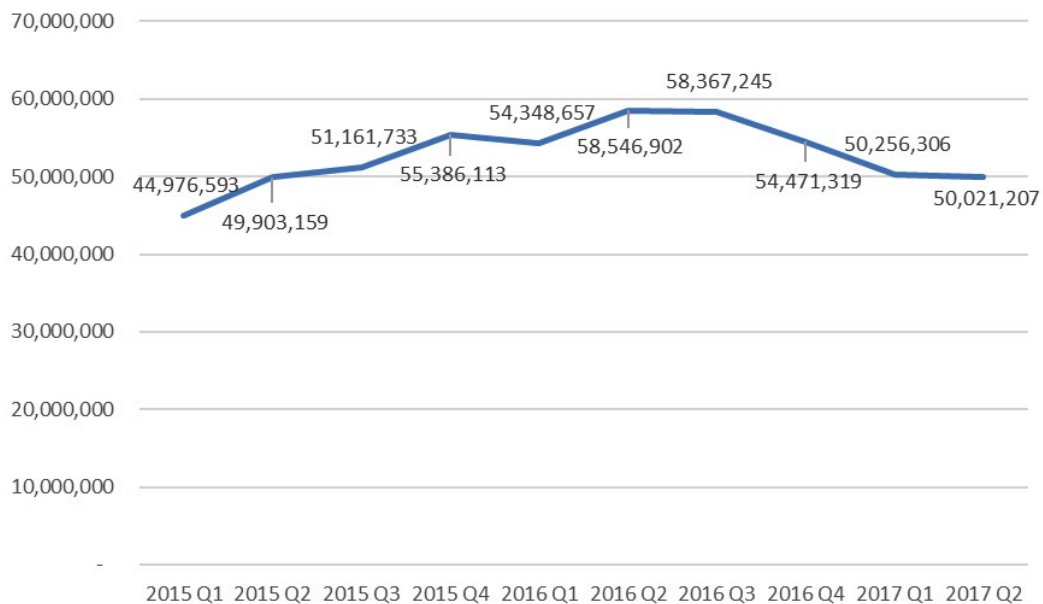


Figure 4. Electronic health record (EHR) volumes downloaded by hospitals and clinics from 2015 to 2017.

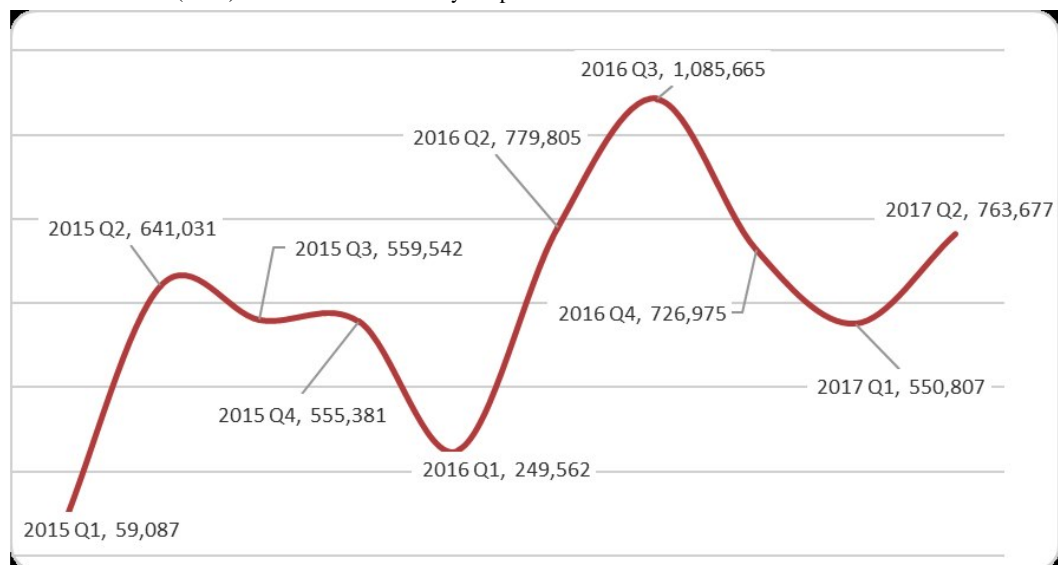


Table 1. Uploaded and downloaded volumes of electronic health records for different accreditation levels, National Health Insurance branches, and electronic health record forms.

Variable and category	Mean (SD)	F value (degrees of freedom)	P value	Scheffe test
Uploaded EHRs^a				
Accreditation level				
Medical centers (A); District hospitals (B); Regional hospitals (C)	20,092,761.7 (1,169,657.3); 22,319,938.9 (2,300,920.9); 10,378,021.1 (1,073,403.2)	154.8 (2)	<.001	A>C; B>A, C
NHI^b branch				
Taipei (A); Northern (B); Central (C); Southern (D); Kaoping (E); Eastern (F)	18,031,142.3 (1,771,219); 6,444,815.4 (636,379.5); 11,366,431.6 (60,1067.2); 7,453,093.0 (708,056.1); 7,898,723.0 (790,655.7); 1,615,391.7 (125,642.2)	360.5 (5)	<.001	A>B, C, D, E, F; B>F; C>B, D, E, F; D>B, F; E>B, D
EHR form				
Laboratory tests (A); Discharge summaries (B); Medical images (C); Outpatient records (D)	26,980,425.5 (2,053,721.2); 675,093.5 (81,408.80); 3,406,489.6 (302,759.2); 21,747,588.4 (1,853,264.3)	887.2 (3)	<.001	A>B, C, D; C>B; D>B, C
Downloaded EHRs				
Accreditation level				
Medical centers (A); Regional hospitals (B); District hospitals (C); Clinics (D)	318,717.8 (332,577.7); 2,523.4 (1623.5); 2,286.3 (996.2); 269,082.1 (199,866.6)	7.63 (3)	<.001	A>B, C, D; B>C; D>B, C
NHI branch				
Taipei (A); Northern (B); Central (C); Southern (D); Kaoping (E); Eastern (F)	130,172.5 (62,431.2); 69,089.4 (37,708.3); 305,230.0 (272,800.0); 58,500.3 (44,413.6); 15,463.7 (12,255.4); 14,173.4 (8116.1)	8.82 (5)	<.001	A>B, D, E, F; B>D, E, F; C>A, B, D, E, F; D>E, F, F
EHR form				
Laboratory tests (A); Discharge summaries (B); Medical images (C); Outpatient records (D)	242,572.3 (168,448.7); 10,081.2 (4778.7); 42,038.8 (19,924.8); 297,937.0 (132,678.9)	17.72 (3)	<.001	A>B, C; C>B; D>A, B, C

^aEHRs: electronic health records.

^bNHI: National Health Insurance.

Variables Affecting the Downloaded Volumes of Electronic Health Records

Table 1 shows that the highest average volume was 318,717.8 downloads by medical centers, followed by clinics with 269,082.1, and the statistical test among hospital accreditation levels was significant ($F=7.63$; $P<.001$). In a comparison of EHR downloaded volumes among the 6 NHI branches, the central NHI branch had the highest volume with 305,230

downloads, and the eastern branch had the lowest volume with 14,173.4 downloads, and it was also statistically significant among the 6 NHI branches ($F=8.82$; $P<.001$). In addition, among downloaded volumes of the 4 EHR forms, outpatient medical records was the highest with 297,937.0 downloads, followed by laboratory tests with 242,572.3, which were both much higher than those for medical images and discharge summaries, and the statistical test was also significant ($F=17.72$; $P<.001$).

Furthermore, we focused on the average download times of the 6 medical images. Table 2 shows that x-rays took the longest at 32.05 seconds, and electrocardiogram (EKG) took 9.92 seconds, which was the shortest, but there was no statistically significant difference. Finally, when each district was compared with the other 5 districts according to the geographical location

in Taiwan, all the highest download volumes occurred within that district as illustrated in Table 3. This indicates that patients tended to obtain treatments within their daily living sphere instead of crossing into a different area, possibly to avoid the extra cost and inconvenience of transportation.

Table 2. Required download times for different medical images. *F* value (degrees of freedom)=0.29 (5); *P*=.91.

Medical images	Mean (SD)
Ultrasound	21.3 (10.4)
X-ray	32.1 (42.8)
Computed tomography	20.3 (7.9)
Magnetic resonance image	31.3 (35.4)
Endoscopy	23.6 (13.2)
Electrocardiogram	9.9 (9.9)

Table 3. Download volumes of electronic health records among six different districts in Taiwan.

Downloaded by	Districts					
	Taipei	Northern	Central	Southern	Kaoping	Eastern
Taipei	1,093,087	169,168	23,674	18,851	9793	6428
Northern	95,138	562,343	31,288	4856	4248	2509
Central	12,773	9415	787,002	23,384	4865	889
Southern	8154	4188	19,085	540,814	15,266	549
Kaoping	1503	1861	1194	5329	146,179	278
Eastern	4671	1640	655	353	1025	133,664

Discussion

Upload Volumes of Electronic Health Records Among Hospitals Tended to Be Steady

The vision of Taiwan's ECC is that hospitals' and clinics' physicians can view patients' most recent 6 months of medical records with the patient's consent and the physician's authorization using the patient's NHI IC card and the physician's professional IC card. Hospitals' upload volumes of EHRs were about 4.5 million to 60 million quarterly (average 52,790,721) during the study period from January 2015 to June 2017. In the beginning, we estimated that upload volumes would slowly decrease as the government's subsidy ended. However, hospital upload volumes of EHRs were quite stable. A possible reason is that the EHR project management office [16], an official EHR exchange standards task force in Taiwan in charge of EHR training, education, demonstration, awards, and subsidies for hospitals and clinics, continued monitoring hospitals' compliance by publishing their weekly upload volumes. Therefore, hospitals had to conduct EHR uploads as a daily operation, and many hospitals used the application programming interface (API) [12] to automatically meet the requirements of the EEC. Figure 3 illustrates that hospital uploads were well maintained and tended to be steady.

Moreover, based on the service capacity and investment of hospital information systems, it is understandable that the total

upload volumes of EHRs by medical centers (>500 beds) and regional hospitals (>300 beds) were significantly higher than those by district hospitals (>100 beds).

Download Volumes of Electronic Health Records Among Hospitals and Clinics Fluctuated

The average quarterly volume of EHR downloads by hospitals and clinics was 592,629 records. Compared with the upload volume listed above, the ratio was about 81:1. In other words, when hospitals uploaded 81 EHRs, only 1 was downloaded by physicians for medical purposes. Under the current policy of the NHI, although the integrated medical care and referral system was strongly promoted and encouraged, the download volumes of EHRs among hospitals and clinics fluctuated. On the basis of physicians' practice behaviors and hospital accreditation, Table 1 shows that medical centers (>500 beds) and clinics downloaded 318,717.8 and 269,082 records, respectively, which accounted for more than 95% of the total download volume. This means that more efforts are needed to implement 2-way referrals among hospitals and clinics for patients' continuous care. Moreover, Table 3 illustrates that each district's internal download volume was significantly higher than that of the other 5 different districts. We believe that patients were constrained by time and costs, so seeking cross-regional medical treatment is rare in Taiwan, and this also proved a resident life circle exists in local people.

In Taiwan, in addition to EHR exchanges provided by the EEC, the NHI administration also launched a Pharma Cloud system in 2015 [17]. Physicians in hospitals and clinics were requested to inquire into patients' medications during previous visits throughout this system. Under the single-payer NHI system, hospitals and clinics have to comply with regulations of the NHI administration, or otherwise, their reimbursements will be denied. However, download volumes of EHRs did not decline because of the NHI administration's service. On the contrary, as Figure 4 illustrates, volumes significantly increased in 2016 because of the API for the EEC. It facilitated hospitals automatically downloading EHRs a day before a patient's visit. Especially in the third quarter of 2016 (Figure 4), the download volume surged because medical centers put a lot of effort into EHR exchanges to meet the requirements of hospital accreditation.

We believe that effective health information exchange can avoid repeated examinations and medications [18,19] and reduce unnecessary medical expenses [6,20]. Thus, the CMS of the United States requires health care providers who accept government subsidies to achieve the criteria of meaningful use [9]. Although the volume of EHRs exchanged by hospitals and clinics gradually increased in Taiwan, whether it really contributes to the continuous care of patients is still unclear and worthy of further study.

Download Times of the 4 Electronic Health Record Forms Required by Hospitals and Clinics Did Not Significantly Differ

Table 2 illustrates that download times for ultrasound, EKGs, computed tomography, magnetic resonance imaging, and endoscopic images, and x-rays were between 9.92 and 32.05 seconds. When hospitals and clinics downloaded different types of medical images, the required times did not significantly differ ($P > .914$). However, when we interviewed some physicians, they complained that the waiting times for downloads were too long, especially for x-rays, and it definitely affected their intention to use this service. The gap between our data and physicians' experiences may come from the time measurement. We counted the time from when the EEC received a request from a hospital to when the gateway at the hospital received the downloaded medical image. Due to high outpatient service volumes and a variety of physician practice styles, it was difficult for us to detect how much time elapsed from the hospital gateway to the physician's terminal when these medical images were being checked. Moreover, in some hospitals, the computing priority of the information system may have affected download speeds, so physicians were unable to promptly review downloaded images.

Building Up an Integrated Cross-Hospital Health Care Model

Data exchanges among health organizations are based on interoperable EHRs. The services provided by EEC met the requirements of functional and semantic interoperability. We

adopted the HL7/CDA R2 and Digital Imaging and Communications in Medicine standards to generate clinical documents and IHE XDS profile for the communication infrastructure. Therefore, a physician can read laboratory test reports, medical images, discharge summaries, and outpatient medical records through these interoperability standards by EHR exchanges. Taiwan has invested more than 10 years in building up the EHR infrastructure and legal system [21]. The final goal of the government is to provide continuous care for patients and achieve cross-hospital integrated health care. Many similar examples in the United States and Europe have actively implemented cross-hospital integrated health care. Regional Health Information Exchange Organizations in the United States, which are in charge of cross-regional data exchanges, have focused on the meaningful use in various health care settings. Kaiser Permanente Health Connect [22] provides continuous care to 1.1 million insured persons through EHRs from its 38 hospitals and 650 clinics and gives patients the latest medical information. SmartCare [23] in Europe has developed a standard interoperable platform to share data with 23 regional stakeholders to provide integrated health care services. A LinkCare project called NEXES, an integrated health care shared knowledge community, supports healthy and independent living by chronically ill patients and the elderly by health care professionals. It is an ICT-enabled integrated care program in large-scale trials (5200 patients), targeting prevalent chronic conditions (mainly chronic obstructive pulmonary disease, chronic heart failure, and diabetes) and is being run in Barcelona and Alicante, Spain, and Athens, Greece, in Europe [24]. We believe that the interoperability and connectivity of EHRs are not only the future of health care but they are also big challenges around the globe. Thus, the government has to play a leading role in defining policies and offering incentives to encourage health care organizations to engage in cross-hospital integrated health care.

Limitations

Due to high outpatient service volumes and a variety of physician practice styles, we could not detect how much time it took from hospital gateways to physicians' terminals when medical data were being checked. We suggest that future studies measure the time physicians actually spend on data review except on data only from the ECC.

Conclusions

After years of operation, volumes of EHRs uploaded to the EEC by hospitals were stable but download volumes of EHRs fluctuated. The primary goal of the EEC for promoting cross-hospital data exchanges was achieved. However, the meaningful use of EHRs among hospitals and clinics needs further encouragement and promotion for reducing unnecessary medications and examinations and enhance the quality of care. We suggest that the government's leading role and collective collaboration with health care organizations are important keys to providing effective health information exchanges.

Acknowledgments

The authors would like to thank the Ministry of Health and Welfare and EBM Technologies for providing the data to support this study.

Conflicts of Interest

None declared.

References

1. Wang SJ, Middleton B, Prosser LA, Bardon CG, Spurr CD, Carchidi PJ, et al. A cost-benefit analysis of electronic medical records in primary care. *Am J Med* 2003 Apr 1;114(5):397-403. [doi: [10.1016/S0002-9343\(03\)00057-3](https://doi.org/10.1016/S0002-9343(03)00057-3)] [Medline: [12714130](https://pubmed.ncbi.nlm.nih.gov/12714130/)]
2. Impact Advisors. 2015 Aug 26. Realizing value from an enterprise EHR Investment-Results from a survey of CHIME members URL: <https://www.impact-advisors.com/document/realizing-value-from-an-enterprise-ehr-investment/> [accessed 2019-02-11] [WebCite Cache ID 7663IDNNy]
3. Marsolo K, Margolis PA, Forrest CB, Colletti RB, Hutton JJ. A digital architecture for a network-based learning health system: integrating chronic care management, quality improvement, and research. *EGEMS (Wash DC)* 2015;3(1):1168 [FREE Full text] [doi: [10.13063/2327-9214.1168](https://doi.org/10.13063/2327-9214.1168)] [Medline: [26357665](https://pubmed.ncbi.nlm.nih.gov/26357665/)]
4. Frame A, LaMantia M, Bynagari BB, Dexter P, Boustani M. Development and implementation of an electronic decision support to manage the health of a high-risk population: the enhanced EMR aging brain care software. *EGEMS (Wash DC)* 2013;1(1):1009 [FREE Full text] [doi: [10.13063/2327-9214.1009](https://doi.org/10.13063/2327-9214.1009)] [Medline: [25848560](https://pubmed.ncbi.nlm.nih.gov/25848560/)]
5. Wasser EJ, Prevedello LM, Sodickson A, Mar W, Khorasani R. Impact of a real-time computerized duplicate alert system on the utilization of computed tomography. *JAMA Intern Med* 2013 Jun 10;173(11):1024-1026 [FREE Full text] [doi: [10.1001/jamainternmed.2013.543](https://doi.org/10.1001/jamainternmed.2013.543)] [Medline: [23609029](https://pubmed.ncbi.nlm.nih.gov/23609029/)]
6. Eftekhari SY, Singh R, Gopal R. Search eLibrary: SSRN. Do Health Information Exchanges Deter Repetition of Medical Services? URL: <https://ssrn.com/abstract=2971503> [WebCite Cache ID 2971503]
7. O'Donnell HP, Patel V, Kern LM, Barrón Y, Teixeira P, Dhopeswarkar R, et al. Healthcare consumers' attitudes towards physician and personal use of health information exchange. *J Gen Intern Med* 2011 Sep;26(9):1019-1026 [FREE Full text] [doi: [10.1007/s11606-011-1733-6](https://doi.org/10.1007/s11606-011-1733-6)] [Medline: [21584839](https://pubmed.ncbi.nlm.nih.gov/21584839/)]
8. Gellert GW, Webster S, Gillean J, Melnick E, Kanzaria H. Should US doctors embrace electronic health records? *Br Med J* 2017 Dec 24;356:j242. [doi: [10.1136/bmj.j242](https://doi.org/10.1136/bmj.j242)] [Medline: [28119282](https://pubmed.ncbi.nlm.nih.gov/28119282/)]
9. SearchHealthIT. What is meaningful use stage 2? URL: <https://searchhealthit.techtarget.com/definition/meaningful-use-stage-2> [WebCite Cache ID 1234]
10. Ministry of Health and Welfare - Electronic Medical Record Exchange Center. URL: <https://eec.mohw.gov.tw/?PAGE=Overview> [WebCite Cache ID 1234]
11. HIMSS Analytics. EMRAM-Stage 6 Hospitals URL: <http://www.himssanalyticsasia.org/emram/stage6hospitals.asp> [WebCite Cache ID 1234]
12. Taichung Veterans General Hospital. URL: http://www.vghtc.gov.tw/GipOpenWeb/wSite/lp?ctNode=55378&mp=1231&idPath=55377_55378&nowPage=1&pagesize=15 [WebCite Cache ID 1234]
13. Yan YL, Lu CL. An assessment of the application of pharma cloud system to the National Health Insurance Program of Taiwan and the result in hospitals. *Stud Health Technol Inform* 2016;225:959-960. [Medline: [27332431](https://pubmed.ncbi.nlm.nih.gov/27332431/)]
14. YouTube. Administration NHI URL: <https://www.youtube.com/watch?v=4udxv13hQMo> [WebCite Cache ID 1234]
15. Huang SP, Pan YT, Chen MS. My Health Bank 2.0-making a patron saint for people's health. *J Formos Med Assoc* 2017 Feb;116(2):69-71 [FREE Full text] [doi: [10.1016/j.jfma.2016.05.008](https://doi.org/10.1016/j.jfma.2016.05.008)] [Medline: [27421173](https://pubmed.ncbi.nlm.nih.gov/27421173/)]
16. Ministry of Health and Welfare. EHRs implementation and promotion website of Department of Health and Welfare URL: <http://emr.mohw.gov.tw/introduction.aspx> [WebCite Cache ID 1234]
17. Ministry of Health and Welfare: National Health Insurance Administration. PharmaCloud system URL: https://www1.nhi.gov.tw/Nhi_E-LibraryPubWeb/PharmaCloud/PharmaCloud.aspx?FType=11 [WebCite Cache ID 1234]
18. Dimitropoulos LP, Patel V, Scheffler SA, Posnack S. Public attitudes toward health information exchange: perceived benefits and concerns. *Am J Manag Care* 2011 Dec;17(12 Spec No):SP111-SP116 [FREE Full text] [Medline: [22216769](https://pubmed.ncbi.nlm.nih.gov/22216769/)]
19. Patel VD, Dhopeswarkar RV, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. *Inform Health Soc Care* 2011 Dec;36(4):233-252. [doi: [10.3109/17538157.2011.554930](https://doi.org/10.3109/17538157.2011.554930)] [Medline: [21851182](https://pubmed.ncbi.nlm.nih.gov/21851182/)]
20. Yaraghi N. An empirical analysis of the financial benefits of health information exchange in emergency departments. *J Am Med Inform Assoc* 2015 Nov;22(6):1169-1172. [doi: [10.1093/jamia/ocv068](https://doi.org/10.1093/jamia/ocv068)] [Medline: [26117143](https://pubmed.ncbi.nlm.nih.gov/26117143/)]
21. Li YY, Yen JC, Chiu WT, Jian WS, Syed-Abdul S, Hsu MH. Building a national electronic medical record exchange system - experiences in Taiwan. *Comput Methods Programs Biomed* 2015 Aug;121(1):14-20. [doi: [10.1016/j.cmpb.2015.04.013](https://doi.org/10.1016/j.cmpb.2015.04.013)] [Medline: [26001420](https://pubmed.ncbi.nlm.nih.gov/26001420/)]
22. Kaiser Permanente. Electronic Health Records URL: <https://share.kaiserpermanente.org/category/electronic-health-records/> [WebCite Cache ID 1234]

23. SmartCare. URL: <http://www.pilotsmartcare.eu/home.html> [WebCite Cache ID 1234]
24. Cano I, Alonso A, Hernandez C, Burgos F, Barberan-Garcia A, Roldan J, et al. An adaptive case management system to support integrated care services: lessons learned from the NEXES project. J Biomed Inform 2015 Jun;55:11-22 [FREE Full text] [doi: [10.1016/j.jbi.2015.02.011](https://doi.org/10.1016/j.jbi.2015.02.011)] [Medline: [25796455](https://pubmed.ncbi.nlm.nih.gov/25796455/)]

Abbreviations

API: application programming interface
CDA R2: Clinical Document Architecture, Release 2
CMS: Centers for Medicare and Medicaid Services
EEC: Electronic Medical Record Exchange Center
EHR: electronic health record
EKG: electrocardiogram
EMR: electronic medical record
EP: eligible professional
HL7: Health Level 7
IC: integrated circuit
ICT: information and communications technology
IHE: Integrating the Healthcare Enterprise
NHI: National Health Insurance
TVGH: Taichung Veterans General Hospital
XDS: Cross-enterprise Document Sharing

Edited by G Eysenbach; submitted 29.10.18; peer-reviewed by YW Huang, M Bestek; comments to author 05.12.18; revised version received 08.01.19; accepted 20.01.19; published 28.03.19.

Please cite as:

Wen HC, Chang WP, Hsu MH, Ho CH, Chu CM

An Assessment of the Interoperability of Electronic Health Record Exchanges Among Hospitals and Clinics in Taiwan

JMIR Med Inform 2019;7(1):e12630

URL: <http://medinform.jmir.org/2019/1/e12630/>

doi: [10.2196/12630](https://doi.org/10.2196/12630)

PMID: [30920376](https://pubmed.ncbi.nlm.nih.gov/30920376/)

©Hsyien-Chia Wen, Wei-Pin Chang, Min-Huei Hsu, Cheng-Hsun Ho, Chi-Ming Chu. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 28.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

SNOMED CT Concept Hierarchies for Computable Clinical Phenotypes From Electronic Health Record Data: Comparison of Intensional Versus Extensional Value Sets

Ling Chu¹, MD; Vaishnavi Kannan¹, MS; Mujeeb A Basit¹, MD, MMSc; Diane J Schaefflein¹, BS, MT(ASCP); Adolfo R Ortuzar¹, BS; Jimmie F Glorioso¹, MS; Joel R Buchanan², MD; Duwayne L Willett¹, MD, MS

¹University of Texas Southwestern Medical Center, Dallas, TX, United States

²University of Wisconsin School of Medicine and Public Health, Madison, WI, United States

Corresponding Author:

Ling Chu, MD

University of Texas Southwestern Medical Center

5323 Harry Hines Boulevard

Dallas, TX, 75390

United States

Phone: 1 214 648 1303

Email: Ling.Chu@UTSouthwestern.edu

Related Article:

This is a corrected version. See correction statement: <https://medinform.jmir.org/2019/3/e14654/>

Abstract

Background: Defining clinical phenotypes from electronic health record (EHR)-derived data proves crucial for clinical decision support, population health endeavors, and translational research. EHR diagnoses now commonly draw from a finely grained clinical terminology—either native SNOMED CT or a vendor-supplied terminology mapped to SNOMED CT concepts as the standard for EHR interoperability. Accordingly, electronic clinical quality measures (eCQMs) increasingly define clinical phenotypes with SNOMED CT value sets. The work of creating and maintaining list-based value sets proves daunting, as does insuring that their contents accurately represent the clinically intended condition.

Objective: The goal of the research was to compare an intensional (concept hierarchy-based) versus extensional (list-based) value set approach to defining clinical phenotypes using SNOMED CT-encoded data from EHRs by evaluating value set conciseness, time to create, and completeness.

Methods: Starting from published Centers for Medicare and Medicaid Services (CMS) high-priority eCQMs, we selected 10 clinical conditions referenced by those eCQMs. For each, the published SNOMED CT list-based (extensional) value set was downloaded from the Value Set Authority Center (VSAC). Ten corresponding SNOMED CT hierarchy-based intensional value sets for the same conditions were identified within our EHR. From each hierarchy-based intensional value set, an exactly equivalent full extensional value set was derived enumerating all included descendant SNOMED CT concepts. Comparisons were then made between (1) VSAC-downloaded list-based (extensional) value sets, (2) corresponding hierarchy-based intensional value sets for the same conditions, and (3) derived list-based (extensional) value sets exactly equivalent to the hierarchy-based intensional value sets. Value set conciseness was assessed by the number of SNOMED CT concepts needed for definition. Time to construct the value sets for local use was measured. Value set completeness was assessed by comparing contents of the downloaded extensional versus intensional value sets. Two measures of content completeness were made: for individual SNOMED CT concepts and for the mapped diagnosis clinical terms available for selection within the EHR by clinicians.

Results: The 10 hierarchy-based intensional value sets proved far simpler and faster to construct than exactly equivalent derived extensional value set lists, requiring a median 3 versus 78 concepts to define and 5 versus 37 minutes to build. The hierarchy-based intensional value sets also proved more complete: in comparison, the 10 downloaded 2018 extensional value sets contained a median of just 35% of the intensional value sets' SNOMED CT concepts and 65% of mapped EHR clinical terms.

Conclusions: In the EHR era, defining conditions preferentially should employ SNOMED CT concept hierarchy-based (intensional) value sets rather than extensional lists. By doing so, clinical guideline and eCQM authors can more readily engage specialists in vetting condition subtypes to include and exclude, and streamline broad EHR implementation of condition-specific decision support promoting guideline adherence for patient benefit.

(*JMIR Med Inform* 2019;7(1):e11487) doi:[10.2196/11487](https://doi.org/10.2196/11487)

KEYWORDS

SNOMED CT; value sets; clinical phenotypes; population health; pragmatic clinical study

Introduction

Overview

Given widespread adoption of electronic health records (EHRs) certified to follow terminology standards, why does achieving interoperable clinical phenotype definitions remain challenging? Practical approaches to analytic interoperability among EHR-originated datasets would provide value both for population health analytics and clinical research [1-3]. Clinical registries define most patient subpopulations—important clinical phenotypes—by either a shared condition or a shared exposure (eg, to a type of procedure or medication) [4]. EHRs now encode patient conditions in clinical terminologies mapped to SNOMED CT, an international comprehensive clinical terminology [5-7]. By federal standard, exchanging patient conditions (problems) between EHRs via health information exchanges employs SNOMED CT concepts.

Accordingly, clinical quality measures derived from EHR data increasingly define clinical phenotypes with SNOMED CT concept value sets, analogous to the International Classification of Diseases (ICD) code value sets traditionally defined for claims data. Initial SNOMED CT value sets primarily have taken an “extensional” form—that is, an enumerated list of terms—in keeping with the long-standing structure of ICD code value sets [8]. But SNOMED CT, being a polyhierarchical ontology, affords the powerful option of employing rule-based or “intensional” value sets leveraging the relationships within the ontology. Such intensional value sets can more concisely identify included and excluded subtypes of a clinical condition by referring to SNOMED CT’s hierarchical “is a” supertype-subtype (parent-child) relationships. Those subtype relationships can be a close match to clinicians’ thinking about clinical phenotypes and the subtypes of conditions they wish to be included or excluded. In a report on 125 such hierarchy-based value sets, we’ve shown they also are simple to create in an EHR and employ in an analytic data warehouse [9].

In the United States, the governmental Centers for Medicare and Medicaid Services (CMS) employs public quality measures to help assure the quality of health care for Medicare beneficiaries, primarily the elderly or disabled. In this study, we examined value sets defining 10 conditions referenced by 2018 Centers for Medicare and Medicaid Services (CMS) high-priority electronic clinical quality measures (eCQMs) for adults. We compare corresponding intensional versus extensional SNOMED CT value sets for their conciseness, time to construct, and completeness of SNOMED CT concept inclusion. We also compare their completeness in covering the

SNOMED CT-mapped clinical terms selectable by clinicians within the EHR as patient Problem List entries and encounter diagnoses, since those selections ultimately drive clinical phenotypes for population health activities and pragmatic clinical studies employing EHR source data.

Extensional SNOMED CT Value Sets

Why Value Sets?

Transactional source data from administrative and clinical information systems typically include diagnosis information encoded in either ICD or SNOMED CT. Value sets of specified ICD or SNOMED CT terms define conditions (clinical phenotypes) for use in clinical guidelines, clinical quality measures, and patient registries [4]. Two categories of conditions commonly need to be defined: (1) one or more primary, population-defining conditions and (2) comorbid conditions used for exclusions and/or risk stratification.

Why SNOMED CT Value Sets?

To be certified for the Meaningful Use program in the United States, EHRs must be able to transmit patient diagnosis information to another EHR using SNOMED CT–encoded concepts [10,11]. Thus, in most EHRs clinicians enter patient conditions onto their Problem List by using either SNOMED CT directly, or, more commonly, a clinician-friendly clinical terminology premapped to SNOMED CT concepts. Both methods enable preserving a higher level of clinical fidelity and relevant clinical detail than ICD does due to the enhanced clinical specificity of SNOMED CT [9].

Why Extensional SNOMED CT Value Sets?

Extensional value sets refer to simple lists of codes or concepts. ICD value sets traditionally have been constructed this way, in keeping with the structure of ICD [12,13]. As the need for SNOMED CT value sets arose, the same approach was continued. The innovative Value Set Authoring Tool made available in 2013 by the Value Set Authority Center (VSAC) initially supported only creation of extensional value sets [14].

Challenges with Extensional SNOMED CT Value Sets

Extensional value sets, as specifically enumerated lists, are brittle and prone to “break” or become stale with updates to the underlying terminology. SNOMED CT updates can include addition of new clinical concepts or refining an existing concept by creating or expanding its “descendant” concepts. Preexisting extensional value set lists cannot handle these automatically and may require frequent reupdating after new SNOMED CT version releases, followed by reimportation or copying into every EHR or other system employing the value set.

Some extensional value sets include many items, which inhibits rapid human comprehension of exactly which subtypes of a given clinical condition are being included and excluded. Thus clinical vetting of such value sets becomes laborious. Similarly, construction of the value set and performing quality assurance are correspondingly difficult and labor-intensive. Inaccuracies in value sets can significantly affect clinical quality measure calculations [15].

Intensional Value Sets of SNOMED CT Concept Hierarchies

SNOMED CT intensional value sets, by contrast, are rule-based and leverage the polyhierarchy structure of SNOMED CT. That is, one can include or exclude an entire “tree” of real-world condition subtypes via a single reference to a SNOMED CT concept and all its descendants. Combining such tree references with simple Boolean logic (or with SNOMED CT Expression Constraint Language) enables efficient definition of a desired clinical phenotype [9]. For instance, osteoporosis and all of its subtypes can be defined by reference to one SNOMED CT concept (SCT ID 64859006 Osteoporosis) and all its descendants. The corresponding extensional list would require 42 SNOMED CT concepts to fully define. In turn, in our EHR 2287 diagnosis clinical terms map to this single SNOMED CT concept hierarchy; a clinician selecting any one of these for a patient’s diagnosis would automatically include them in the broad computable clinical phenotype of osteoporosis.

Possible benefits of SNOMED CT intensional value sets include closely matching how clinicians think about what condition subtypes to include or exclude from a given clinical phenotype. Being able to reference the entire tree of a concept’s descendants enables far simpler, succinct value set definitions that are easier to understand and construct. Additionally, they should be more resilient to change and less likely to omit descendants and break with future SNOMED CT concept additions. Consequently, intensional value sets have potential to be simultaneously simpler and more complete and thus more useful for population health analytics and clinical research using EHR data.

Objective of the Study

For each of 10 conditions (clinical phenotypes), evaluate the differences between an intensional (concept hierarchy-based) versus extensional (list-based) SNOMED CT value set approach in (1) conciseness, (2) time to create, and (3) completeness of both SNOMED CT concepts included and relevant clinical terms available for clinician selection in an EHR.

Methods

Selection of Value Sets

Value sets included in this study were identified starting from the CMS website for choosing Merit-Based Incentive Payment System (MIPS) quality measures [16]. MIPS measures were filtered for high-priority measures and data submission method of EHR, yielding 21 candidate measures. Four measures covering the following 4 common adult conditions were selected: hypertension, diabetes mellitus, depression, and prostate cancer.

Next, the online VSAC “search value sets” feature was employed to find condition-defining SNOMED CT value sets for these measures [17]. Value sets were first filtered for CMS eCQM Release = “eCQM Update 2018 EP-EC and EH” and Code System = “SNOMEDCT.” Then each of the eCQMs was selected individually, displaying the related SNOMED CT value sets. Any value sets specifying a condition (diagnosis) were included, yielding an initial total of 12 SNOMED CT extensional value sets (see [Multimedia Appendix 1](#)).

Software

Creation of EHR vendor-neutral SNOMED CT intensional value sets and automatic derivation of extensional value sets were both done using Symedical (Clinical Architecture LLC), a clinical terminology management and mapping software tool for health care professionals. SNOMED CT intensional value sets (groupers) for EHR-based registry and clinical decision support functionality were created using the grouper management features of our EHR, Epic (Epic Systems Corporation). The clinical terminology vocabulary within University of Texas Southwestern Medical Center’s Epic EHR during this study was the proprietary IMO Problem IT terminology, version 2018 R1 (Intelligent Medical Objects Inc), mapped to the SNOMED CT International Edition July 2017 release and the SNOMED CT US Edition September 2017 release.

Procedures

Using the VSAC website’s “export value set results” feature, the list of codes for each SNOMED CT extensional value set was exported to Excel (Microsoft Corp) for subsequent comparison.

Comparable intensional (rule-based) value set diagnosis groupers for these conditions were established in our EHR. The majority already existed, having been created for disease registries and/or clinical decision support [1]; two were newly created for this study (pain related to prostate cancer, personality disorder). Identically matching intensional value sets were then constructed in Symedical (in addition to Epic), and the time to construct each intensional value set recorded. Intensional value sets were defined using a “search, drill-up, drill-down” approach previously described [9]. Existing and newly-defined intensional value sets were vetted by medical informaticians and clinicians by deriving the full list of included SNOMED CT concepts for review.

To enable meaningful direct comparison with intensional value sets, two combinations of VSAC value sets were performed prior to comparing the SNOMED CT concept lists: (1) chronic kidney disease, stage 5, (CKD-5) was combined with end-stage renal disease (ESRD) since clinically they refer to the same condition, and so only one intensional value set covered both, and (2) major depression including remission was combined with dysthymia, as together they constitute the condition of depressive disorders covered by a single intensional value set. This yielded a final set of 10 clinical conditions for comparison. The eCQMs, VSAC value set identifiers, and extensional value set contents are available in Excel format in [Multimedia Appendix 1](#).

The pregnancy value set (2.16.840.1.113883.3.526.3.378) downloaded from VSAC was found to include concepts focused on pregnancy itself but in general did not include concepts for complications or disorders of pregnancy. Our existing EHR-based intensional value set for pregnancy deliberately included the latter to provide a broad net for identifying any pregnant patients via EHR-entered diagnoses and problems. Accordingly, to better match the VSAC contents, we constructed a second narrow intensional value set for pregnancy based on the pregnancy conditions listed in the VSAC extensional value set by deliberately omitting SNOMED CT concepts for pregnancy-related conditions (eg, complication occurring during pregnancy, disorder of pregnancy). The VSAC extensional pregnancy value set was compared separately with both the broad and the narrow intensional pregnancy value sets.

For each intensional value set, a corresponding extensional value set list was automatically derived using Symedical (ie, a list of all included SNOMED CT concept descendants). These derived extensional value sets were downloaded and stored for subsequent analysis. The intensional value sets and corresponding derived extensional value sets are available in Excel format in [Multimedia Appendix 2](#).

Measures and Outcomes

Value Set Definition Conciseness

Conciseness of value set definition was measured simply by the number of SNOMED CT concepts needed to fully define the set, either as a list (extensional value set) or the number of concepts in the defining rule (intensional value set). A dimensionless ratio to define was calculated in two forms:

Ratio to define (download) = (# concepts in VSAC-downloaded extensional value set) / (# concepts in intensional value set defining rule)

Ratio to define (derived) = (# concepts in derived extensional value set) / (# concepts in intensional value set defining rule)

Time to Construct

The purpose of the “Time to Construct” measure is to gauge the time needed at each healthcare organization to construct in their local systems, such as their EHR, an approved value set definition received from a defining group such as VSAC (or a local clinical terminology committee). The preceding upfront “time to define” the value set, including iterative clinical review, is purposefully not included.

The time to construct in Symedical each of 11 intensional value sets (including both pregnancy value set versions) as well as 3 of the extensional value sets (CKD-5 & ESRD; prostate cancer; pain related to prostate cancer) was measured. From this a best-fit linear equation was derived: $\text{time (min)} = 0.4177 * (\# \text{ SNOMED CT concepts}) + 3.8707$. This corresponds to an obligate time of just under 4 minutes to construct any value set (eg, for configuring basic common settings), plus approximately 0.42 minutes (25 seconds) to add each SNOMED CT concept. The time to construct the remaining extensional value sets was estimated using this equation.

The difference in time to construct an extensional versus an intensional value set was calculated as (time to construct extensional value set) – (time to construct intensional value set), expressed in minutes. The dimensionless ratio was calculated as (time to construct extensional value set) / (time to construct intensional value set).

Completeness: SNOMED CT Concepts

For each of 10 conditions, the list of SNOMED CT concepts included in the VSAC-downloaded set and the intensional-derived set were compared. The total number of concept discrepancies present in one set and not the other was assessed by summing two discrepancy types:

- Number of concepts present in the VSAC-downloaded set but not in the intensional-derived set
- Number of concepts present in the intensional-derived set but not in the VSAC-downloaded set

Since virtually all of the SNOMED CT concepts in the downloaded extensional value sets were included in the corresponding intensional-derived value set, the ratio of the two was calculated as: (# concepts in intensional-derived set) / (# concepts in VSAC-downloaded set), expressed as a number greater than 1. The percentage of SNOMED CT concepts included in the downloaded extensional value set was calculated as: (# concepts in VSAC-downloaded set) / (# concepts in intensional-derived set), expressed as a percentage.

Completeness: Electronic Health Record Clinical Term Coverage

To evaluate the impact of condition-specific discrepancies, value sets were created in the EHR in both an intensional form (existing) and an extensional form (to exactly match the VSAC list of concepts, without including descendants). The EHR automatically creates a compiled list of IMO-sourced clinical terms mapped to the SNOMED CT value set. These IMO clinical terms comprise the diagnoses visible to clinicians for selection as Problem List entries and as diagnoses to associate with patient orders, encounters, and professional charges. The number of clinical terms compiled for each intensional and extensional value set was recorded. Comparisons were then performed on the number of clinical terms available for selection by clinicians in the EHR that would result in patient inclusion in a given clinical phenotype.

Just as for SNOMED CT concept completeness, the ratio of the two was calculated as (# clinical terms from intensional-derived set) / (# clinical terms from VSAC-downloaded set), expressed as a number greater than or equal to 1. The percentage of clinical terms covered by the downloaded extensional value set was calculated as (# clinical terms from VSAC-downloaded set) / (# clinical terms from intensional-derived set), expressed as a percentage.

Results

Overall Format of Result Tables

Tabulated comparisons by each of the 10 conditions follow. Summary calculated measures are included at the bottom of each table. In addition to overall sums and ratios, the median

of the 10 condition-specific values was selected as the primary measure of central tendency. This method was chosen a priori to avoid the potential for skew if one or more conditions exhibited marked difference from the others or contained many more concepts. The minimum, maximum, and range across the 10 conditions are also reported.

For pregnancy, both the narrow and broad definitions are shown in tabular form; however, only the more narrow intensional value set based on the CMS extensional value set was used in all summary calculations (to avoid double-counting). Use of the narrow pregnancy definition reduces the reported differences between intensional and extensional value sets so that the

summary findings and conclusions shown are conservative. Were the broad pregnancy definition selected instead, the magnitude of effects would be larger. All tables are available in Excel format in [Multimedia Appendix 3](#).

Value Set Conciseness

We expected that intensional value sets should be more concise to construct by leveraging the hierarchical supertype-subtype structure of SNOMED CT. [Table 1](#) shows that the median number of SNOMED CT concepts employed to define a condition with the VSAC value sets was 21.5 concepts versus only 3.0 for intensional value sets.

Table 1. Clinical phenotypes with value set definition conciseness and time to construct.

Condition name	SNOMED CT concepts to define			Time to construct					
	Extensional ^a (n)	Intensional (n)	Ratio to define (download)	Extensional ^b (n)	Ratio to define (derived)	Extensional ^b (min)	Intensional (min)	Diff min	Ratio (ext/int) ^c
CKD-5 ^d and ESRD ^e	5	3	1.7	27	9	17	3	12	5
Hypertension	12	3	4	233	77.7	101	3	81	27.9
Kidney transplant	8	8	1	36	4.5	19	15	3	1.2
Pregnancy, narrow	35	6	5.8	156	26	69	4	54	14.5
Pregnancy, broad ^f	35	20	1.8	1262	63.1	531	12	415	35.6
Bipolar disorder	99	2	49.5	109	54.5	49	5	37	8.5
Depression and dysthymia	72	2	36	151	75.5	67	5	51	11.3
Personality disorders	26	2	13	48	24	24	5	17	4.4
Diabetes	89	6	14.8	441	73.5	188	6	147	25.5
Prostate cancer	12	3	4	19	6.3	10	7	5	1.8
Pain related to prostate cancer	17	14	1.2	20	1.4	12	7	6	1.8
Overall sum, ratio, or %	375	49	7.5	1240	25.3	556	60	413	7.9
Set of 10 conditions									
Overall median	21.5	3	4.9	78.5	25	36.7	5	27.2	6.7
Overall minimum	5	2	1	19	1.4	10	3	3	1.2
Overall maximum	99	14	49.5	441	77.7	188	15	147	27.9
Overall range	94	12	48.5	422	76.2	178	12	144	26.7

^aDownload from Value Set Authority Center.

^bDerived from intensional.

^cExt/int: extensional/intensional

^dCKD-5: chronic kidney disease, stage 5.

^eESRD: end-stage renal disease.

^fNot included in summary calculated measures (overall sum, median, minimum, maximum, range).

For the full extensional value sets derived from the intensional rules, the median number of concepts-to-define was 78.5 concepts. The median ratio of concepts needed to fully define an equivalent extensional value set was 25 times that needed for the intensional value set rule.

As one example, the clinical phenotype of personality disorder is specified by 26 SNOMED CT concepts in the downloaded

extensional value set ([Figure 1](#)). In contrast, the corresponding intensional value set rule (inferring intent of subtypes desired from examining the VSAC downloaded list) includes just two concepts: (1) Personality disorder (SCT ID 33449004), including descendants, AND NOT (2) Organic personality disorder (SCT ID 36217008), including descendants. This rule includes all 26 SNOMED CT concepts in the VSAC extensional value set plus an additional 22 closely related concepts that reasonably belong,

for 48 included concepts and a concept ratio-to-define of 48/2 or 24.0.

Time to Construct

Not surprisingly, more concise value sets are easier and faster to construct, perform quality assurance on, review, and update as needed. As shown in Table 1, it takes 6 to 8 times longer to construct an extensional value set completely equivalent in contents to an intensional value set (median 6.7, overall ratio 7.9). In this set, constructing intensional value sets (groupers) for all 10 conditions was accomplished in just 1 hour (60 minutes) of keyboard time, while constructing the equivalent

extensional value sets required nearly 11 hours (650 minutes). The median construction time for these 10 conditions was 5 minutes for an intensional value set and 37 minutes for an equivalent extensional value set.

Completeness: SNOMED CT Concepts

SNOMED CT is updated twice yearly [18] and an intensional rule-based approach presumably should be more resilient to updates by automatically including new descendants within an existing included hierarchy, for instance. Accordingly, we examined the relative completeness of downloaded extensional versus corresponding intensional value sets.

Figure 1. (a) SNOMED CT extensional value set list (26 items) downloaded for the condition personality disorder, shown as green-colored members of a SNOMED CT hierarchy (partial diagram only). Red-colored items aren't on the list, downloaded from the Value Set Authority Center. (b) Matched intensional value set combining SNOMED CT hierarchies with Boolean logic: personality disorder (disorder; 33449004), including descendants AND NOT organic personality disorder (disorder; 36217008), including descendants. (c) Implementation of intensional value set in an electronic health record (EHR) (Epic Systems). (d) Implementation of intensional value set in an EHR-agnostic terminology software program (Symedical). Also shown is part of the exactly equivalent extensional value set (containing 48 SNOMED CT concepts), automatically derived from the intensional logic.

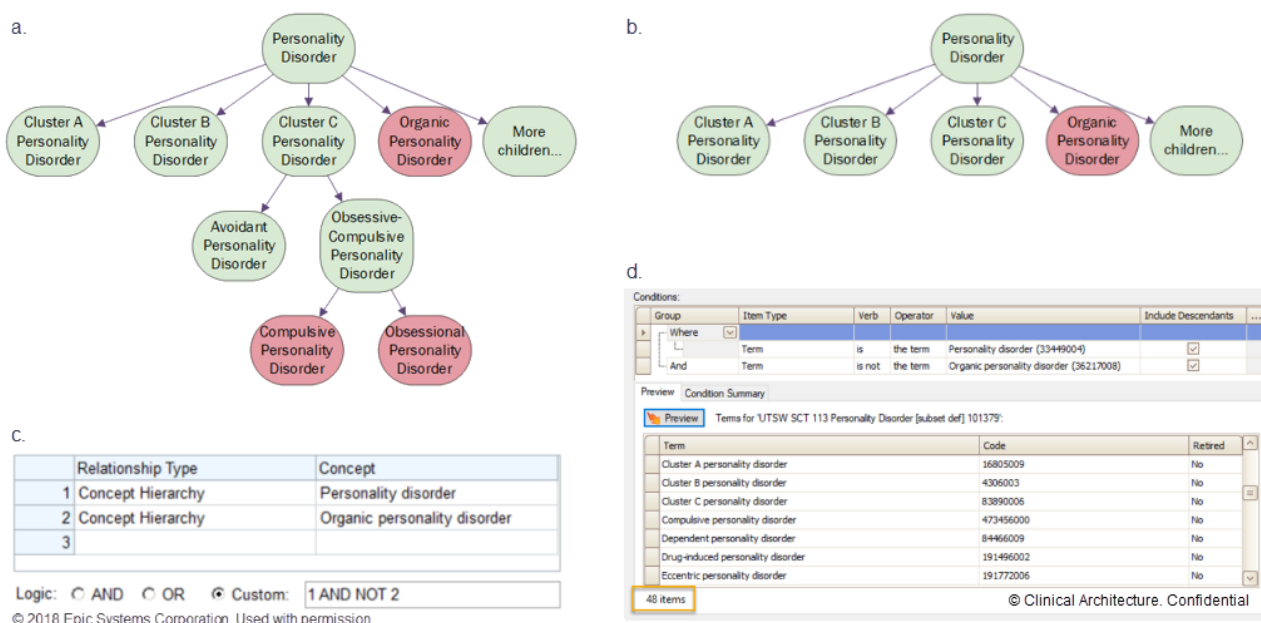


Table 2 compares the number of SNOMED CT concepts included in the full extensional list derived from the intensional rule versus the extensional 2018 list downloaded from VSAC. Across the 10 conditions, the full derived list included a median of 3.3 times as many SNOMED CT concepts as the corresponding downloaded list (range 1.1 to 19.4). In percentage terms, a median of only 35% of SNOMED CT concepts in the full derived extensional list were present in the corresponding downloaded extensional list (range 5% to 91%), as shown in Figure 2 (left panel).

The vast majority of discrepancies between the two sets of extensional lists (877/889, 98.7%) were present in the intensional-derived list only and missing from the VSAC-downloaded list (Table 2); 1.3% (12/889) of concepts in the VSAC download were not in the intensional-derived list. Of these 12, 6 were kidney transplant procedural concepts rather than disorder or condition concepts and had no corresponding diagnosis clinical terms defined in the EHR's clinical terminology. The remaining 6 were judged clinically relevant omissions from the intensional-derived list.

Completeness: Coverage of Relevant Electronic Health Record Clinical Terms

Pragmatic clinical trials, registries, and other research projects that rely on EHR data for clinical phenotypes need the most accurate and complete value sets possible to define primary and comorbid conditions. We thus compared the number of EHR clinical terms (sourced from IMO, overall n>800,000) selectable by clinicians that are in extensional versus intensional value set compiled lists (see Table 3 and Figure 2, right panel).

In 9 of 10 conditions, the number of EHR clinical terms identified using the downloaded extensional value set was less than when using the corresponding intensional value set, in some cases dramatically so. In this subset of 10 conditions, a median 65% of the EHR diagnostic clinical terms selectable by clinicians in a commonly used EHR are included when using a published list-based extensional value set compared with using a corresponding concept hierarchy-based intensional value set. That is, a median of 35% of clinician-selectable diagnosis terms in the EHR for defining a clinical phenotype are missing when using a 2018 downloaded extensional value set.

Table 2. Comparison of downloaded versus derived SNOMED CT value set contents.

Condition name	SNOMED CT concepts included in value set				Extensional concept discrepancies		
	Extensional ^a (n)	Extensional ^b (n)	Ratio (derived/ VSAC)	SNOMED CT concepts includ- ed in VSAC download (%)	Total (n)	Only in VSAC download (n)	Only in inten- sional-derived (n)
CKD-5 ^c and ESRD ^d	5	27	5.4	19	22	0	22
Hypertension	12	233	19.4	5	223	1	222
Kidney transplant	8	36	4.5	22	40	6	34
Pregnancy, narrow	35	156	4.5	22	121	0	121
Pregnancy, broad ^e	35	1262	36.1	3	1227	0	1227
Bipolar disorder	99	109	1.1	91	10	0	10
Depression and dysthymia	72	151	2.1	48	79	0	79
Personality disorders	26	48	1.8	54	22	0	22
Diabetes	89	441	5	20	358	3	355
Prostate cancer	12	19	1.6	63	11	2	9
Pain related to prostate cancer	17	20	1.2	85	3	0	3
Overall sum, ratio, or %	375	1240	3.3	30	889	12	877
Set of 10 conditions							
Overall median	21.5	78.5	3.3	35	31	0	28
Overall minimum	5	19	1.1	5	3	0	3
Overall maximum	99	441	19.4	91	358	6	355
Overall range	94	422	18.3	86	355	6	352

^aDownload from Value Set Authority Center (VSAC).

^bDerived from intensional.

^cCKD-5: chronic kidney disease, stage 5.

^dESRD: end-stage renal disease.

^eNot included in summary calculated measures (overall sum, median, minimum, maximum, range).

Figure 2. Left: SNOMED CT concepts included in 2018 extensional value sets as a percentage of those in intensional value sets. Right: Electronic health records clinical terms included using the 2018 extensional value sets as a percentage of those using intensional value sets. CKD-5: chronic kidney disease, stage 5; ESRD: end-stage renal disease; Ca: cancer.

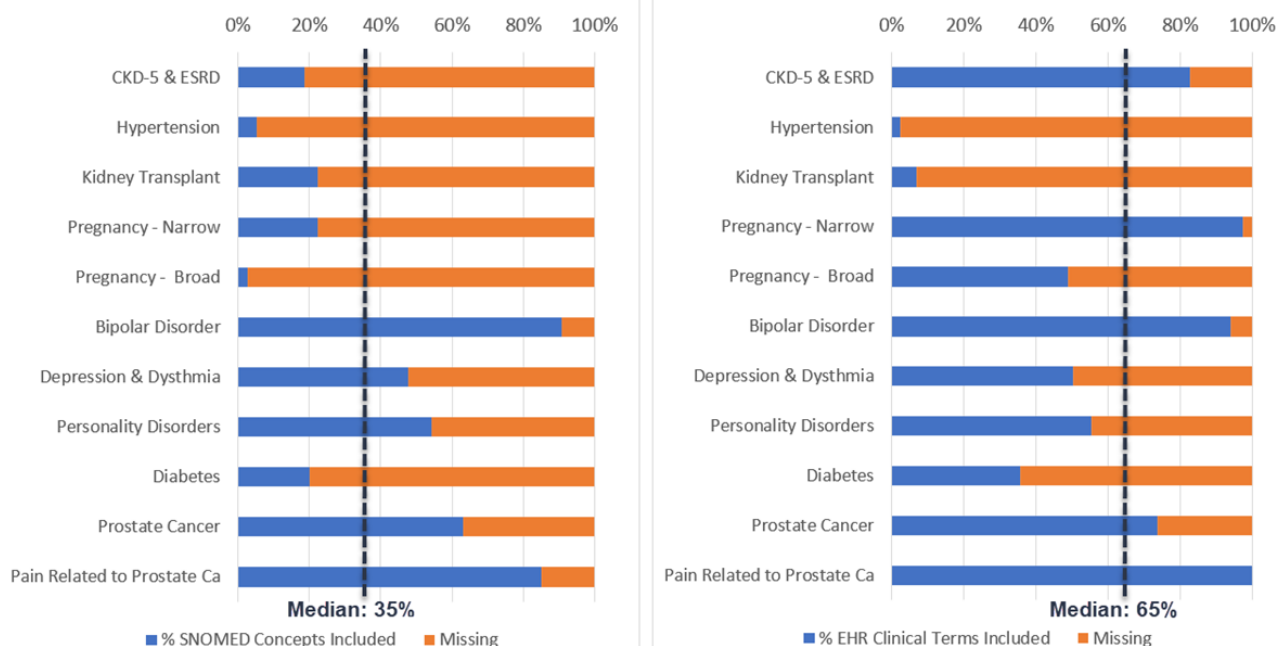


Table 3. Number of diagnosis clinical terms selectable by clinicians in the electronic health record by source of value set.

Condition name	Extensional ^a (n)	Intensional (n)	Ratio (intensional/VSAC)	EHR ^b clinical terms included in VSAC download (%)
CKD-5 ^c and ESRD ^d	485	586	1.2	83
Hypertension	131	5473	41.8	2
Kidney transplant	49	708	14.4	7
Pregnancy, narrow	23,429	24,043	1	97
Pregnancy, broad ^e	23,429	47,812	2	49
Bipolar disorder	1640	1744	1.1	94
Depression and dysthymia	978	1946	2	50
Personality disorders	401	724	1.8	55
Diabetes	11,997	33,707	2.8	36
Pain related to prostate cancer	149	149	1	100
Overall sum, ratio, or %	39,445	69,332	1.8	57
Set of 10 Conditions				
Overall median	443	1234	1.6	65
Overall minimum	49	149	1	2
Overall maximum	23,429	33,707	41.8	100
Overall range	23,380	33,558	40.8	98

^aDownload from Value Set Authority Center (VSAC).

^bEHR: electronic health record.

^cCKD-5: chronic kidney disease, stage 5.

^dESRD: end-stage renal disease.

^eNot included in summary calculated measures (overall sum, median, minimum, maximum, range).

Discussion

Principal Results

For 10 conditions referenced by the CMS 2018 high-priority clinical quality measures, we compared extensional SNOMED CT lists of codes downloaded in the fall of 2018 from the VSAC with intensional (rule-based) value sets for the same conditions. Intensional value set definitions were far more concise (median number of concepts needed for equivalent value sets 3 vs 75), faster to construct (median 5 vs 37 minutes each), and more complete. VSAC-downloaded value sets were missing a median 65% of the SNOMED CT concepts included in the intensional rule-based value sets and 35% of the mapped diagnosis clinical terms selectable by clinicians within the EHR.

The conciseness of intensional value sets expedites construction in the EHR. This should also streamline vetting with busy clinical experts and harmonizing multiple value set specifications of the same real-world condition. Many systems will directly import large extensional value set files, mitigating the value set construction time/workload difference for those organizations. Still, someone must create the source value sets initially and periodically update them, and some customer organizations will have to enter them manually as well. For those, the large time reduction offered by intensional value set definitions remains an advantage. Because intensional value sets are rule-based and can include references to all descendants of a parent concept, they are more likely to include all relevant concepts than an enumerated list. That is, they are less likely to inadvertently omit descendant concepts and more likely to remain complete following future SNOMED CT updates.

Clinical Phenotyping for Clinical-Translational Studies Using Electronic Health Record Data

Pragmatic clinical trials, registries, and other clinical and translational research studies employing EHR data for computable clinical phenotypes (rather than manual abstraction) rely on having as accurate and complete value sets to define primary and comorbid conditions as feasible [19,20]. Concern typically arises about missing diagnosis data not yet entered in the EHR by clinicians on the patient's Problem List or as Encounter Diagnoses. While Problem List completeness in particular remains a subject of active inquiry and improvement efforts [21-26], this study raises a different concern for completeness of value set definitions when physicians and advanced practice providers have conscientiously recorded their patients' specific diagnoses in the EHR. In this subset of 10 conditions, a median 35% of the EHR diagnostic clinical terms selectable by clinicians in a commonly used EHR are missing when using a published extensional value set compared with using a simpler rule-based intensional value set. Patients for whom those missing EHR terms are selected by clinicians will fail to be included in the selected population with the clinical phenotype. Controlling for comorbid conditions in multivariable modeling will similarly be negatively impacted by missing clinical EHR terms. Defining clinical phenotypes more completely with rule-based intensional value sets leveraging SNOMED CT's hierarchical structure advances the feasibility and reliability of pragmatic clinical studies and learning health

care system cycles conducted with EHR data produced during normal clinical care [27-29].

Analytic Interoperability for Population Health

With the expansion of clinically integrated networks and cross-institution specialty registries to provide and measure value-based care, definition of subpopulations of patients becomes crucial for risk assessment and tailored interventions [1,30-32]. Many networks encompass a variety of EHRs. Since the designated interoperability language between EHRs for diagnoses (conditions) is SNOMED CT, employing SNOMED CT value sets enables EHR-agnostic consistent definition of subpopulations for registries, clinical decision support to promote best practices within the EHR, care gap closure, and quality measurement [9]. This provides analytic interoperability across disparate EHRs even if using clinical terminologies from different vendors. The populations that would most benefit from intervention may change over time, thus generating requests for new computable clinical phenotype definitions. The conciseness and clinical understandability of intensional value sets streamline rapid-cycle definition and vetting by specialists, as well as more facile and consistent implementation across a broad range of EHRs, population health tools, and clinical settings. These advantages make employing SNOMED CT concept rule-based intensional value sets a higher quality, better fit-for-purpose method for defining computable clinical phenotypes for population health than traditional extensional lists.

Authoring Practice Guidelines and Electronic Clinical Quality Measures for Streamlined Implementation

With the expansion in medical knowledge and appreciation of the complexities of achieving optimal care for subpopulations of patients with a wide variety of conditions, the number of clinical practice guidelines continues to grow [33,34]. Significant effort and expense (in terms of experts' time) goes into writing consensus guidelines and optimal practices for a condition. Achieving real-world practice change takes a long time and is often incomplete [35-37]. EHR-based clinical decision support has been shown to improve clinical process measures across multiple clinical domains [38-48].

Yet current guidelines can be difficult to implement as point-of-care clinical decision support to help "make the right thing the easy thing to do" for busy clinicians within their daily work tool, the EHR [49-53]. Non-value-added work can include:

- Translating prose definition of conditions covered by guideline, conditions excluded, and comorbid conditions into value sets implementable in EHRs to cover clinical terms/codes present in EHRs in practice
- Translating prose definitions of medication types and/or procedure types into EHR-implementable value sets
- Translating prose descriptions (and some flow charts and decision trees, if constructed ambiguously) into implementable decision algorithms for clinical decision support logic [54]

EHRs have local codes that can hamper implementation, but increasingly these are mapped to standard terminology codes

to achieve interoperability with other EHRs as organizations participate in health information exchanges [27].

To accelerate implementation, we propose that specialty guideline and eCQM writing committees include a medical informaticist (as either a consultant or a formal member of the writing group representing a clinical informatics specialty society). During initial guideline development discussion and through subsequent detailed specification, the medical informaticist could then assist specialist experts on the committee in expressing the clinical conditions relevant to the guideline or eCQM in a SNOMED CT supertype-subtype form, readily implementable in an EHR or other internet-accessible repository as a concise, easily shareable intensional rule (Figure 3).

Doing so would avoid the considerable extra work of constructing a de novo extensional value set, vetting the full list with clinical experts, distributing it, and having multiple teams of EHR analysts and clinical informaticists around the country independently reverse-engineer the list into a supertype-subtype rule-based form to gain its benefits of conciseness, maintainability, completeness, and understandability for their local EHR implementation. In lean terms, that extra work (red arrows in Figure 3) could be considered non-value-adding waste. In contrast, coproducing concise, shareable SNOMED CT intensional value sets contemporaneously with the guideline and/or eCQM specification would expedite practical dissemination of clinical decision support to promote the new best practice in a consistent subpopulation of patients across the country, matching the guideline writing specialists' intent.

Limitations

Challenges When Using SNOMED CT

Navigating the SNOMED CT Hierarchy and Selecting Concepts for an Intensional Value Set

Because of the polyhierarchical structure of SNOMED CT, potential exists for inadvertently including descendant branches and/or individual concepts which do not belong. The “search, drill-up, drill-down” approach employed mitigates that risk by explicitly exploring if the currently-selected concept in a SNOMED CT hierarchy browser is too general or too narrow [9]. A helpful additional mitigation strategy is to expand the intensional rule to show all included SNOMED CT concepts as a derived extensional list (we used Symedical for this purpose), then having a clinician view this list for any additional concepts which should be excluded. These then similarly can

be evaluated with the “search, drill-up, drill-down” method to find the optimal concept in the hierarchy for exclusion along with its descendants.

Changes to SNOMED CT

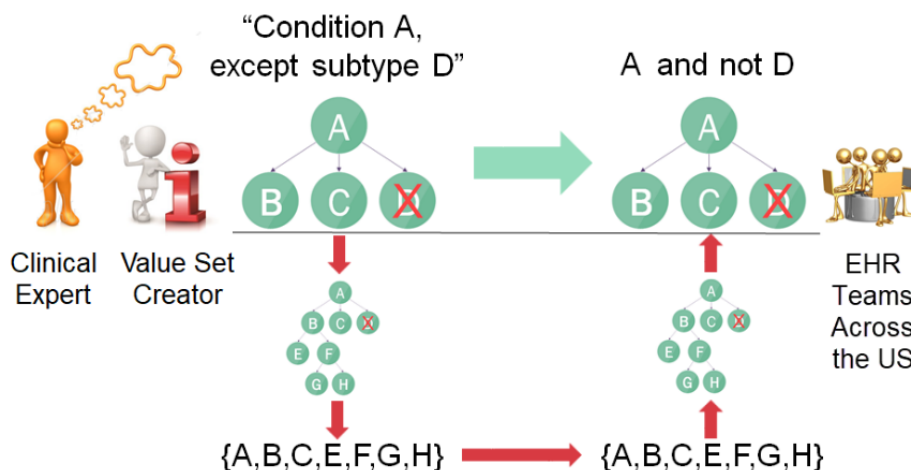
Importantly, although intensional value sets retain accuracy and completeness across many updates to SNOMED CT's contents, they are not impervious to changes [55-57]. With intensional rules referencing SNOMED CT's hierarchical structure, additions of new descendants are generally automatically included. Some value sets may need clinical vetting for updates after new SNOMED CT releases, perhaps particularly when an intensional rule includes some, but not all, of a parent concept's children. To enhance rapid re-vetting when needed, automated detection of new SNOMED CT concept additions that are within the span of a given rule-based grouper would be useful. One question to explore further is whether a specific inclusion strategy (include these specific siblings) versus an exclusion strategy (include all the children of the parent except these specific children) proves more resilient (remains more complete and accurate).

Migrations of existing SNOMED CT concepts to a different location in the hierarchy due to cleanup of SNOMED CT quality issues [58] pose a different challenge, although in many cases an intensional value set will handle the correction gracefully [9]. As clinicians and medical informaticists work with intensional value sets to define important clinical phenotypes, iterative improvements in SNOMED CT's hierarchical arrangement will likely ensue, following the data quality aphorism “what gets used, gets better” [1].

Scope of This Paper's Analysis and Differences in Value Set Intent

One limitation of this paper is that the comparable intensional value sets were developed and vetted only at one institution (University of Texas Southwestern Medical Center) and cannot be considered to represent national specialty society views. However, our experience demonstrates feasibility for a medical informaticist to build an initial candidate rule for defining a condition, then identify any clinical inclusion/exclusion questions for vetting with a clinician specializing in the condition [9]. For multi-institutional and/or specialty society vetting, a Modified Delphi technique can be employed as was successfully used by Buchanan [59] previously to gain working consensus across institutions. Our vision is that increasingly intensional value sets are produced as a byproduct of clinical guideline and eCQM authoring, dramatically reducing the need for individual institutions to reinvent the wheel (Figure 3).

Figure 3. Clinicians' thinking about type/subtype inclusion criteria matches SNOMED CT concept hierarchy implementation in electronic health records (EHRs). Deriving flat code lists (red arrows) requires reverse engineering by each EHR team to achieve the benefits of a concept hierarchy-based definition.



This study only covers 10 conditions and may not be representative of all and so should be considered merely as a deep dive into one set of conditions for CMS-designated eCQMs (as described in the Methods section). We took a conservative approach in matching intensional definitions to VSAC-downloaded extensional definitions, otherwise differences reported between intensional and extensional value set completeness would have been even greater. Specifically, for pregnancy we did not use our existing broad pregnancy intensional value set: instead we constructed a new, much more narrowly defined value set intended to match the scope of the VSAC-released pregnancy value set. Similarly, for pain related to prostate cancer, we lacked an existing intensional value set and constructed our new intensional value set closely mirroring the contents of the VSAC extensional value set. Both result in minimizing differences between the extensional and intensional approaches.

On the other hand, for hypertension our existing intensional value set includes all forms of hypertension (meant to represent the scope covered by recent hypertension guidelines [60-62]), whereas the VSAC-downloaded extensional value set was specific to essential (primary) hypertension. The latter did not include SNOMED CT concepts for the general concept of “Hypertensive disorder, systemic arterial (disorder)” not specified to be primary or secondary, for secondary hypertension, or for “Complication of systemic hypertensive disorder (disorder)”. Replacing our existing hypertension intensional value set with one mirroring the contents of the VSAC-downloaded essential hypertension value set would have increased this condition’s values for % completeness of both SNOMED CT concepts and EHR terms. However, our pre-specified use of medians instead of means (averages) results in no change in the overall median values reported of 35% completeness for SNOMED CT concepts and 65% completeness for EHR clinical terms.

Inconsistencies in SNOMED CT polyhierarchy “is a” definitions may lead to inadvertent inclusion of unwanted descendants of a seemingly wholly-appropriate SNOMED CT concept. Use of the “search, drill-up, drill-down” method during intensional value set definition can reduce the likelihood of this, as can clinical review of the full list of included SNOMED CT concepts derived from the intensional definition [9]. As discovered, such unwanted descendants can be specifically excluded in the intensional rule. Also requests to update the “is a” relationship in SNOMED CT to a more specific parent(s) can be made through the SNOMED CT Content Request Service. Once the subsumption has been updated in SNOMED CT, the value set intensional rule typically can be further simplified.

Conclusions

Although extensional lists of codes have long been used for ICD-based value sets, the use of extensional lists of SNOMED CT codes is suboptimal and fails to leverage the capabilities and clinical relevance of ontological relationships within SNOMED CT. Compared with SNOMED CT extensional (list) value sets, intensional (rule-based) value sets are far simpler to create, maintain, understand, and vet with specialist clinicians. For the 10 conditions studied here from the 2018 CMS high-value eCQMs for the MIPS program, intensional value sets also proved substantially more complete than their corresponding extensional list versions: a median 35% of diagnosis terms selectable in the EHR by clinicians were missing when using a downloaded extensional value set, with risk of failing to identify patients with a given clinical phenotype despite physician-entered discrete diagnoses in the EHR.

Consequently, in the EHR era we believe defining conditions as computable clinical phenotypes preferentially should employ SNOMED CT concept hierarchy-based (intensional) value sets rather than extensional lists. By doing so, clinical guideline and eCQM authors can streamline broad EHR implementation of condition-specific decision support promoting guideline adherence and patient benefit.

Acknowledgments

Research reported in this publication was supported by the University of Texas Southwestern Health System Information Resources department and by the National Center for Advancing Translational Sciences of the National Institutes of Health under award number UL1TR001105. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Source electronic clinical quality measures and corresponding VSAC-downloaded extensional value sets in MS Excel format.

[[XLSX File \(Microsoft Excel File\), 29KB - medinform_v7i1e11487_app1.xlsx](#)]

Multimedia Appendix 2

Intensional value set definitions with corresponding derived extensional value sets in MS Excel format.

[[XLSX File \(Microsoft Excel File\), 129KB - medinform_v7i1e11487_app2.xlsx](#)]

Multimedia Appendix 3

Study data for tables and analyses in MS Excel format.

[[XLSX File \(Microsoft Excel File\), 42KB - medinform_v7i1e11487_app3.xlsx](#)]

References

1. Kannan V, Fish JS, Mutz JM, Carrington AR, Lai K, Davis LS, et al. Rapid development of specialty population registries and quality measures from electronic health record data: an agile framework. *Methods Inf Med* 2017 Jun 14;56(99):e74-e83 [FREE Full text] [doi: [10.3414/ME16-02-0031](#)] [Medline: [28930362](#)]
2. El Fadly A, Rance B, Lucas N, Mead C, Chatellier G, Lastic P, et al. Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform. *J Biomed Inform* 2011 Dec;44 Suppl 1:S94-S102 [FREE Full text] [doi: [10.1016/j.jbi.2011.07.007](#)] [Medline: [21888989](#)]
3. Richesson RL, Smerek MM, Blake CC. A framework to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications. *EGEMS (Wash DC)* 2016;4(3):1232 [FREE Full text] [doi: [10.13063/2327-9214.1232](#)] [Medline: [27563686](#)]
4. Agency for Healthcare Research and Quality. Registries for Evaluating Patient Outcomes: A User's Guide URL: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-guide-3rd-edition_research.pdf [accessed 2018-12-23] [WebCite Cache ID 74sK5qQjo]
5. Spackman KA, Campbell KE, Côté RA. SNOMED RT: a reference terminology for health care. *Proc AMIA Annu Fall Symp* 1997;640-644 [FREE Full text] [Medline: [9357704](#)]
6. SNOMED CT Basics. URL: <https://confluence.ihtsdotools.org/display/DOCSTART/4.+SNOMED+CT+Basics> [accessed 2018-12-19] [WebCite Cache ID 74kuHEBzl]
7. SNOMED CT Concept Model. URL: <https://confluence.ihtsdotools.org/display/DOCSTART/6.+SNOMED+CT+Concept+Model> [accessed 2018-01-15] [WebCite Cache ID 75RjPpbzd]
8. Value Set Authority Center: Datasets. URL: <https://healthdata.gov/dataset/value-set-authority-center> [accessed 2018-12-18] [WebCite Cache ID 74kuNcuQr]
9. Willett DL, Kannan V, Chu L, Buchanan JR, Velasco FT, Clark JD, et al. SNOMED CT concept hierarchies for sharing definitions of clinical conditions using electronic health record data. *Appl Clin Inform* 2018 Jul;9(3):667-682 [FREE Full text] [doi: [10.1055/s-0038-1668090](#)] [Medline: [30157499](#)]
10. Office of the National Coordinator for Health Information Technology (ONC), Department of Health and Human Services (HHS). 2015 edition health information technology (Health IT) certification criteria, 2015 edition base electronic health record (EHR), definition, and ONC health IT certification program modifications. Final rule. *Fed Regist* 2015 Oct 16;80(200):62601-62759 [FREE Full text] [Medline: [26477063](#)]
11. Blumenthal D. Launching HITECH. *N Engl J Med* 2010 Feb 4;362(5):382-385. [doi: [10.1056/NEJMp0912825](#)] [Medline: [20042745](#)]
12. Chute CG. Clinical classification and terminology: some history and current observations. *J Am Med Inform Assoc* 2000;7(3):298-303 [FREE Full text] [Medline: [10833167](#)]

13. Benson T, Grieve G. Principles of Health Interoperability NOMED CT, HL7, and FHIR. Third edition. London: Springer; 2016.
14. Khatipov E, Madden M, Chiang P, Chuang P, Nguyen D, D'Souza I, et al. Creating, maintaining and publishing value sets in the VSAC. 2014 Presented at: AMIA Annu Symp Proc; 2014; Washington p. 1459.
15. D'Amore JD, Li C, McCrary L, Niloff JM, Sittig DF, McCoy AB, et al. Using clinical data standards to measure quality: a new approach. *Appl Clin Inform* 2018 Apr;9(2):422-431. [doi: [10.1055/s-0038-1656548](https://doi.org/10.1055/s-0038-1656548)] [Medline: [29898468](https://pubmed.ncbi.nlm.nih.gov/29898468/)]
16. Merit-based Incentive Payment System (MIPS) quality measures.: Centers for Medicare and Medicaid Services; 2017. URL: <https://qpp.cms.gov/mips/quality-measures> [accessed 2018-12-19] [WebCite Cache ID 74kudhoRZ]
17. Value Set Authority Center (VSAC).: National Library of Medicine URL: <https://vsac.nlm.nih.gov/welcome> [accessed 2018-12-18] [WebCite Cache ID 74kugdBT6]
18. Release schedule and file formats: SNOMED International. URL: <https://confluence.ihtsdotools.org/display/DOCSTART/13.+Release+Schedule+and+File+Formats> [accessed 2018-12-18] [WebCite Cache ID 74kullvcc]
19. Richesson RL, Horvath MM, Rusincovitch SA. Clinical research informatics and electronic health record data. *Yearb Med Inform* 2014 Aug 15;9:215-223 [FREE Full text] [doi: [10.15265/TY-2014-0009](https://doi.org/10.15265/TY-2014-0009)] [Medline: [25123746](https://pubmed.ncbi.nlm.nih.gov/25123746/)]
20. Adamusiak T, Shimoyama N, Shimoyama M. Next generation phenotyping using the unified medical language system. *JMIR Med Inform* 2014 Mar 18;2(1):e5 [FREE Full text] [doi: [10.2196/medinform.3172](https://doi.org/10.2196/medinform.3172)] [Medline: [25601137](https://pubmed.ncbi.nlm.nih.gov/25601137/)]
21. Galanter WL, Hier DB, Jao C, Sarne D. Computerized physician order entry of medications and clinical decision support can improve problem list documentation compliance. *Int J Med Inform* 2010 May;79(5):332-338. [doi: [10.1016/j.ijmedinf.2008.05.005](https://doi.org/10.1016/j.ijmedinf.2008.05.005)] [Medline: [18599342](https://pubmed.ncbi.nlm.nih.gov/18599342/)]
22. Jao C, Hier D, Galanter W. Automating the maintenance of problem list documentation using a clinical decision support system. *AMIA Annu Symp Proc* 2008 Nov 06:989. [Medline: [18998927](https://pubmed.ncbi.nlm.nih.gov/18998927/)]
23. Wright A, Pang J, Febowitz JC, Maloney FL, Wilcox AR, McLoughlin KS, et al. Improving completeness of electronic problem lists through clinical decision support: a randomized, controlled trial. *J Am Med Inform Assoc* 2012;19(4):555-561 [FREE Full text] [doi: [10.1136/amiainjnl-2011-000521](https://doi.org/10.1136/amiainjnl-2011-000521)] [Medline: [22215056](https://pubmed.ncbi.nlm.nih.gov/22215056/)]
24. Wright A, McCoy AB, Hickman TT, Hilaire DS, Borbolla D, Bowes WA, et al. Problem list completeness in electronic health records: a multi-site study and assessment of success factors. *Int J Med Inform* 2015 Oct;84(10):784-790 [FREE Full text] [doi: [10.1016/j.ijmedinf.2015.06.011](https://doi.org/10.1016/j.ijmedinf.2015.06.011)] [Medline: [26228650](https://pubmed.ncbi.nlm.nih.gov/26228650/)]
25. McEvoy D, Gandhi TK, Turchin A, Wright A. Enhancing problem list documentation in electronic health records using two methods: the example of prior splenectomy. *BMJ Qual Saf* 2018 Jan;27(1):40-47. [doi: [10.1136/bmjqs-2017-006707](https://doi.org/10.1136/bmjqs-2017-006707)] [Medline: [28754813](https://pubmed.ncbi.nlm.nih.gov/28754813/)]
26. Simons SM, Cillessen FH, Hazelzet JA. Determinants of a successful problem list to support the implementation of the problem-oriented medical record according to recent literature. *BMC Med Inform Decis Mak* 2016 Dec 02;16:102 [FREE Full text] [doi: [10.1186/s12911-016-0341-0](https://doi.org/10.1186/s12911-016-0341-0)] [Medline: [27485127](https://pubmed.ncbi.nlm.nih.gov/27485127/)]
27. Mo H, Thompson WK, Rasmussen LV, Pacheco JA, Jiang G, Kiefer R, et al. Desiderata for computable representations of electronic health records-driven phenotype algorithms. *J Am Med Inform Assoc* 2015 Nov;22(6):1220-1230 [FREE Full text] [doi: [10.1093/jamia/ocv112](https://doi.org/10.1093/jamia/ocv112)] [Medline: [26342218](https://pubmed.ncbi.nlm.nih.gov/26342218/)]
28. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med* 2010 Nov 10;2(57). [doi: [10.1126/scitranslmed.3001456](https://doi.org/10.1126/scitranslmed.3001456)] [Medline: [21068440](https://pubmed.ncbi.nlm.nih.gov/21068440/)]
29. Richesson RL, Green BB, Laws R, Puro J, Kahn MG, Bauck A, et al. Pragmatic (trial) informatics: a perspective from the NIH Health Care Systems Research Collaboratory. *J Am Med Inform Assoc* 2017 Dec 01;24(5):996-1001. [doi: [10.1093/jamia/ocx016](https://doi.org/10.1093/jamia/ocx016)] [Medline: [28340241](https://pubmed.ncbi.nlm.nih.gov/28340241/)]
30. Tonner C, Schmajuk G, Yazdany J. A new era of quality measurement in rheumatology: electronic clinical quality measures and national registries. *Curr Opin Rheumatol* 2017 Mar;29(2):131-137 [FREE Full text] [doi: [10.1097/BOR.0000000000000364](https://doi.org/10.1097/BOR.0000000000000364)] [Medline: [27941392](https://pubmed.ncbi.nlm.nih.gov/27941392/)]
31. Drummond FJ, Kinnear H, Donnelly C, O'Leary E, O'Brien K, Burns RM, et al. Establishing a population-based patient-reported outcomes study (PROMs) using national cancer registries across two jurisdictions: the Prostate Cancer Treatment, your experience (PiCTure) study. *BMJ Open* 2015 Apr 17;5(4):e006851 [FREE Full text] [doi: [10.1136/bmjopen-2014-006851](https://doi.org/10.1136/bmjopen-2014-006851)] [Medline: [25888474](https://pubmed.ncbi.nlm.nih.gov/25888474/)]
32. Dulai PS, Singh S, Ohno-Machado L, Sandborn WJ. Population health management for inflammatory bowel disease. *Gastroenterology* 2018 Dec;154(1):37-45. [doi: [10.1053/j.gastro.2017.09.052](https://doi.org/10.1053/j.gastro.2017.09.052)] [Medline: [29122544](https://pubmed.ncbi.nlm.nih.gov/29122544/)]
33. US National Library of Medicine. 2018. Summary of professional guidelines: National Center for Biotechnology Information URL: <https://www.ncbi.nlm.nih.gov/medgen/docs/guideline/> [accessed 2018-07-08] [WebCite Cache ID 70IW1e1bW]
34. AHRQ. 2018. National Guideline Clearinghouse: Browse by Clinical Specialty URL: <https://www.guideline.gov/browse/clinical-specialty> [accessed 2018-07-08] [WebCite Cache ID 70IVQd23n]
35. Balas EA, Boren SA. Managing clinical knowledge for health care improvement. *Yearb Med Inform* 2000(1):65-70. [Medline: [27699347](https://pubmed.ncbi.nlm.nih.gov/27699347/)]
36. Grant J, Green L, Mason B. Basic research and health: a reassessment of the scientific basis for the support of biomedical science. *Res Eval* 2003 Dec 01;12(3):217-224. [doi: [10.3152/147154403781776618](https://doi.org/10.3152/147154403781776618)]

37. Westfall JM, Mold J, Fagnan L. Practice-based research—blue highways on the NIH roadmap. *JAMA* 2007 Jan 24;297(4):403-406. [doi: [10.1001/jama.297.4.403](https://doi.org/10.1001/jama.297.4.403)] [Medline: [17244837](https://pubmed.ncbi.nlm.nih.gov/17244837/)]
38. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. *Yearb Med Inform* 2016 Aug 02;Suppl 1:S103-S116. [doi: [10.15265/YYS-2016-s034](https://doi.org/10.15265/YYS-2016-s034)] [Medline: [27488402](https://pubmed.ncbi.nlm.nih.gov/27488402/)]
39. Dunn LK, Gephart SM, Raszewski R, Sousa V, Shehorn LE, Abraham J. Integrative review of clinical decision support for registered nurses in acute care settings. *J Am Med Inform Assoc* 2017 Dec 01;24(2):441-450. [doi: [10.1093/jamia/ocw084](https://doi.org/10.1093/jamia/ocw084)] [Medline: [27330074](https://pubmed.ncbi.nlm.nih.gov/27330074/)]
40. Ali SM, Giordano R, Lakhani S, Walker DM. A review of randomized controlled trials of medical record powered clinical decision support system to improve quality of diabetes care. *Int J Med Inform* 2016 Mar;87:91-100. [doi: [10.1016/j.ijmedinf.2015.12.017](https://doi.org/10.1016/j.ijmedinf.2015.12.017)] [Medline: [26806716](https://pubmed.ncbi.nlm.nih.gov/26806716/)]
41. Wall HK, Wright JS. The role of clinical decision support systems in preventing cardiovascular disease. *Am J Prev Med* 2015 Nov;49(5):e83-e84. [doi: [10.1016/j.amepre.2015.07.011](https://doi.org/10.1016/j.amepre.2015.07.011)] [Medline: [26477808](https://pubmed.ncbi.nlm.nih.gov/26477808/)]
42. Tapper EB, Sengupta N, Lai M, Horowitz G. A decision support tool to reduce overtesting for ceruloplasmin and improve adherence with clinical guidelines. *JAMA Intern Med* 2015 Sep;175(9):1561-1562. [doi: [10.1001/jamainternmed.2015.2062](https://doi.org/10.1001/jamainternmed.2015.2062)] [Medline: [26029881](https://pubmed.ncbi.nlm.nih.gov/26029881/)]
43. Holstiege J, Mathes T, Pieper D. Effects of computer-aided clinical decision support systems in improving antibiotic prescribing by primary care providers: a systematic review. *J Am Med Inform Assoc* 2015 Jan;22(1):236-242 [FREE Full text] [doi: [10.1136/amiajnl-2014-002886](https://doi.org/10.1136/amiajnl-2014-002886)] [Medline: [25125688](https://pubmed.ncbi.nlm.nih.gov/25125688/)]
44. Anchala R, Kaptoge S, Pant H, Di Angelantonio E, Franco OH, Prabhakaran D. Evaluation of effectiveness and cost-effectiveness of a clinical decision support system in managing hypertension in resource constrained primary health care settings: results from a cluster randomized trial. *J Am Heart Assoc* 2015 Jan 05;4(1):e001213 [FREE Full text] [doi: [10.1161/JAHA.114.001213](https://doi.org/10.1161/JAHA.114.001213)] [Medline: [25559011](https://pubmed.ncbi.nlm.nih.gov/25559011/)]
45. Murphy EV. Clinical decision support: effectiveness in improving quality processes and clinical outcomes and factors that may influence success. *Yale J Biol Med* 2014 Jun;87(2):187-197 [FREE Full text] [Medline: [24910564](https://pubmed.ncbi.nlm.nih.gov/24910564/)]
46. Fathima M, Peiris D, Naik-Panvelkar P, Saini B, Armour CL. Effectiveness of computerized clinical decision support systems for asthma and chronic obstructive pulmonary disease in primary care: a systematic review. *BMC Pulm Med* 2014 Dec 02;14:189 [FREE Full text] [doi: [10.1186/1471-2466-14-189](https://doi.org/10.1186/1471-2466-14-189)] [Medline: [25439006](https://pubmed.ncbi.nlm.nih.gov/25439006/)]
47. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012 Jul 3;157(1):29-43. [doi: [10.7326/0003-4819-157-1-201207030-00450](https://doi.org/10.7326/0003-4819-157-1-201207030-00450)] [Medline: [22751758](https://pubmed.ncbi.nlm.nih.gov/22751758/)]
48. Dixon BE, Alzeer AH, Phillips EO, Marrero DG. Integration of provider, pharmacy, and patient-reported data to improve medication adherence for type 2 diabetes: a controlled before-after pilot study. *JMIR Med Inform* 2016 Feb 08;4(1):e4. [doi: [10.2196/medinform.4739](https://doi.org/10.2196/medinform.4739)] [Medline: [26858218](https://pubmed.ncbi.nlm.nih.gov/26858218/)]
49. Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks: seeking to address research issues underlying implementation successes and failures. *J Biomed Inform* 2018 Dec;78:134-143. [doi: [10.1016/j.jbi.2017.12.005](https://doi.org/10.1016/j.jbi.2017.12.005)] [Medline: [29246790](https://pubmed.ncbi.nlm.nih.gov/29246790/)]
50. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10(6):523-530 [FREE Full text] [doi: [10.1197/jamia.M1370](https://doi.org/10.1197/jamia.M1370)] [Medline: [12925543](https://pubmed.ncbi.nlm.nih.gov/12925543/)]
51. Shiffman RN, Wright A. Evidence-based clinical decision support. *Yearb Med Inform* 2013;8(1):120-127. [Medline: [23974559](https://pubmed.ncbi.nlm.nih.gov/23974559/)]
52. Basit MA, Baldwin KL, Kannan V, Flahaven EL, Parks CJ, Ott JM, et al. Agile acceptance test-driven development of clinical decision support advisories: feasibility of using open source software. *JMIR Med Inform* 2018 Apr 13;6(2):e23 [FREE Full text] [doi: [10.2196/medinform.9679](https://doi.org/10.2196/medinform.9679)] [Medline: [29653922](https://pubmed.ncbi.nlm.nih.gov/29653922/)]
53. Lyman JA, Cohn WF, Bloomrosen M, Detmer DE. Clinical decision support: progress and opportunities. *J Am Med Inform Assoc* 2010 Oct;17(5):487-492 [FREE Full text] [doi: [10.1136/jamia.2010.005561](https://doi.org/10.1136/jamia.2010.005561)] [Medline: [20819850](https://pubmed.ncbi.nlm.nih.gov/20819850/)]
54. Boxwala AA, Rocha BH, Maviglia S, Kashyap V, Meltzer S, Kim J, et al. A multi-layered framework for disseminating knowledge for computer-based decision support. *J Am Med Inform Assoc* 2011 Dec;18 Suppl 1:i132-i139 [FREE Full text] [doi: [10.1136/amiajnl-2011-000334](https://doi.org/10.1136/amiajnl-2011-000334)] [Medline: [22052898](https://pubmed.ncbi.nlm.nih.gov/22052898/)]
55. Lee D, Cornet R, Lau F. Implications of SNOMED CT versioning. *Int J Med Inform* 2011;80(6):442-453. [Medline: [21450517](https://pubmed.ncbi.nlm.nih.gov/21450517/)]
56. Wade G. Implementing SNOMED CT for quality reporting: avoiding pitfalls. *Appl Clin Inform* 2011;2(4):534-545. [Medline: [23616894](https://pubmed.ncbi.nlm.nih.gov/23616894/)]
57. Wade G, Rosenbloom S. The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. *J Biomed Inform* 2009;42(3):490-493. [Medline: [19285570](https://pubmed.ncbi.nlm.nih.gov/19285570/)]
58. Bodenreider O, Smith B, Kumar A, Burgun A. Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies. *Artif Intell Med* 2007 Mar;39(3):183-195 [FREE Full text] [doi: [10.1016/j.artmed.2006.12.003](https://doi.org/10.1016/j.artmed.2006.12.003)] [Medline: [17241777](https://pubmed.ncbi.nlm.nih.gov/17241777/)]

59. Buchanan J. Accelerating the benefits of the problem oriented medical record. *Appl Clin Inform* 2017 Dec 15;8(1):180-190 [FREE Full text] [doi: [10.4338/ACI-2016-04-IE-0054](https://doi.org/10.4338/ACI-2016-04-IE-0054)] [Medline: [28197621](https://pubmed.ncbi.nlm.nih.gov/28197621/)]
60. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). *JAMA* 2014 Feb 05;311(5):507-520. [doi: [10.1001/jama.2013.284427](https://doi.org/10.1001/jama.2013.284427)] [Medline: [24352797](https://pubmed.ncbi.nlm.nih.gov/24352797/)]
61. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: [10.1016/j.jacc.2017.11.006](https://doi.org/10.1016/j.jacc.2017.11.006)] [Medline: [29146535](https://pubmed.ncbi.nlm.nih.gov/29146535/)]
62. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH Guidelines for the management of arterial hypertension. *Eur Heart J* 2018 Sep 01;39(33):3021-3104. [doi: [10.1093/eurheartj/ehy339](https://doi.org/10.1093/eurheartj/ehy339)] [Medline: [30165516](https://pubmed.ncbi.nlm.nih.gov/30165516/)]

Abbreviations

CKD-5: chronic kidney disease, stage 5
CMS: Centers for Medicare and Medicaid Services
ESRD: end-stage renal disease
eQM: electronic clinical quality measure
EHR: electronic health record
ICD: International Classification of Diseases
IMO: Intelligent Medical Objects
MIPS: Merit-Based Incentive Payment System
NIH: National Institutes of Health
VSAC: Value Set Authority Center

Edited by G Eysenbach; submitted 15.07.18; peer-reviewed by P Greco, J Fountain, I Mircheva; comments to author 30.08.18; revised version received 23.11.18; accepted 09.12.18; published 16.01.19.

Please cite as:

Chu L, Kannan V, Basit MA, Schaefflein DJ, Ortuzar AR, Glorioso JF, Buchanan JR, Willett DL

SNOMED CT Concept Hierarchies for Computable Clinical Phenotypes From Electronic Health Record Data: Comparison of Intensional Versus Extensional Value Sets

JMIR Med Inform 2019;7(1):e11487

URL: <http://medinform.jmir.org/2019/1/e11487/>

doi: [10.2196/11487](https://doi.org/10.2196/11487)

PMID: [30664458](https://pubmed.ncbi.nlm.nih.gov/30664458/)

©Ling Chu, Vaishnavi Kannan, Mujeeb A Basit, Diane J Schaefflein, Adolfo R Ortuzar, Jimmie F Glorioso, Joel R Buchanan, Duwayne L Willett. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 16.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Utilization of the Veterans Affairs' Transgender E-consultation Program by Health Care Providers: Mixed-Methods Study

John R Blosnich^{1,2*}, PhD, MPH; Keri L Rodriguez^{1,2*}, PhD; Kristina L Hruska^{1*}, MS; Dio Kavalieratos^{1,2*}, PhD; Adam J Gordon^{3,4}, MD, MPH; Alexis Matza^{5*}, PhD; Susan M Mejia^{5,6*}, DBA; Jillian C Shipherd^{5,7,8*}, PhD; Michael R Kauth^{5,6,9*}, PhD

¹Center for Health Equity Research and Promotion, Department of Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA, United States

²Division of General Internal Medicine, Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

³Informatics, Decision-Enhancement, and Analytic Sciences Center, Department of Veterans Affairs Salt Lake City Health Care System, Salt Lake City, UT, United States

⁴Department of Internal Medicine, School of Medicine, University of Utah, Salt Lake City, UT, United States

⁵Lesbian, Gay, Bisexual, and Transgender Health Program, Office of Patient Care Services, Veterans Health Administration, Washington, DC, United States

⁶South Central Mental Illness Research, Education and Clinical Center, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, United States

⁷National Center for Posttraumatic Stress Disorder, Department of Veterans Affairs Boston Healthcare System, Boston, MA, United States

⁸Boston University School of Medicine, Boston, MA, United States

⁹Department of Psychiatry, Baylor College of Medicine, Houston, TX, United States

*these authors contributed equally

Corresponding Author:

John R Blosnich, PhD, MPH

Center for Health Equity Research and Promotion

Department of Veterans Affairs Pittsburgh Healthcare System

University Drive C (151C-U), Building 30

Pittsburgh, PA, 15240-1001

United States

Phone: 1 412 360 2138

Email: john.blosnich@va.gov

Abstract

Background: In 2015, the Department of Veterans Affairs (VA) nationally implemented a transgender e-consultation (e-consult) program with expert clinical guidance for providers.

Objective: This mixed-methods project aimed to describe providers' program experiences, reasons for nonuse of the program, and ways to improve the program use.

Methods: From January to May 2017, 15 urban and rural VA providers who submitted at least one e-consult in the last year participated in semistructured interviews about their program experiences, which were analyzed using content analysis. From November to December 2017, 53 providers who encountered transgender patients but did not utilize the program participated in a brief online survey on the reasons for nonuse of the program and the facilitators encouraging use.

Results: Qualitative analysis showed that providers learned of the program through email; colleagues; the electronic health record (EHR) system; and participation in the VA Lesbian, Gay, Bisexual, and Transgender committees or educational trainings. Providers used the program to establish care plans, hormone therapy recommendations, sexual and reproductive health education, surgical treatment education, patient-provider communication guidance, and second opinions. The facilitators of program use included understandable recommendations, ease of use through the EHR system, and status as the only transgender resource for rural providers. Barriers to use included time constraints, communication-related problems with the e-consult, impractical recommendations for underresourced sites, and misunderstanding of the e-consult purpose. Suggestions for improvement included addition of concise or sectioned responses, expansion of program awareness among providers or patients, designation of a follow-up contact person, and increase in provider education about transgender veterans and related care. Quantitative analysis showed that the common reasons for nonuse of the program were no knowledge of the program (54%), no need of the program

(32%), and receipt of help from a colleague outside of e-consult (24%). Common suggestions to improve the program use in quantitative analyses included provision of more information about where to find e-consult in the chart, guidance on talking with patients about the program, and e-mail announcements to improve provider awareness of the program. Post hoc exploratory analyses showed no differences between urban and rural providers.

Conclusions: The VA transgender e-consult program is useful for providers, but there are several barriers to implementing recommendations, some of which are especially challenging for rural providers. Addressing the identified barriers and enhancing the facilitators may improve program use and quality care for transgender veterans.

(*JMIR Med Inform* 2019;7(1):e11695) doi:[10.2196/11695](https://doi.org/10.2196/11695)

KEYWORDS

teleconsultation; telemedicine; transgender persons; veterans; veteran health

Introduction

Transgender is an umbrella term that encompasses individuals with gender identities (ie, the core sense of self as male, female, both, or neither) that conflict with societal prescriptions of masculinity and femininity associated with sex assigned at birth. Transgender individuals experience numerous health disparities [1] including health risk (eg, victimization [2]), health outcomes (eg, depression and HIV [3]), and barriers to health care access [4], all of which can create complex care needs. In addition, they often require medical care specific to their gender affirmation (ie, gender experience reflecting their gender identity) including mental health services; evaluations for hormone therapy; prescription and monitoring of hormones; gender-affirming surgery and postoperative care; and specialty services such as urology, endocrinology, and speech therapy [5]. In addition to specific services, there are unique needs related to documentation, particularly sex assigned at birth, which health systems often link to other important medical services including laboratory test values, medication dosages, and critical health screens such as breast or prostate examinations. Although awareness of care for transgender patients is growing, health system-level efforts to improve transgender patients' care are limited. There is a paucity of formative research on how providers can collaborate with specialists to obtain expert clinical guidance on transgender-specific care for their transgender patients [6].

Achieving equity in health and health care, irrespective of gender, is a specific objective for the Department of Veterans Affairs (VA) [7]. The VA has a higher rate of transgender individuals than the United States (US) population [8], and the transgender veteran population has increased annually in the VA over the last decade [9]. Recognizing the need to serve transgender veterans, in 2011, the VA issued a national policy on health care for transgender veterans. To increase the VA's capacity to provide high-quality care to transgender veterans, the Lesbian, Gay, Bisexual, and Transgender (LGBT) Health Program in the VA Office of Patient Care Services developed a transgender e-consultation (e-consult) program to provide expert clinical guidance for VA health care providers; the program was implemented nationally in April 2015 and was described previously [10].

Generally, e-consults provide a technology-based platform through which teams of providers with expertise in a specific medical condition or process of care can assist providers in

medical centers where experts may not exist locally [11]. The VA program allows transgender veterans to receive health care from their local providers, with remote support from experts in transgender health care, thus reducing or eliminating the need for veterans to travel to tertiary VA sites for care. The program includes interdisciplinary teams with expertise in transgender health care at three sites: Minneapolis, Tucson, and Loma Linda. Each expert team comprises at least 4 members (eg, primary care physician, psychologist, social worker, nurse, endocrinologist, and pharmacist). Consultations typically address 4 main topics: mental health evaluation and hormone readiness (eg, questions about diagnosis of gender dysphoria and assessment of readiness for gender-affirming surgery and hormone therapy); psychotherapy (eg, gender counseling and gender-informed treatment for comorbid conditions); primary medical questions (eg, addressing expected outcomes, assessing medical risks, and recommended monitoring of gender affirmation pharmacologically and medically); and prescription of hormone therapy (eg, how to safely prescribe hormone therapy for gender affirmation).

Since the e-consult is a relatively new resource for VA providers, there is a critical need for formative evaluation that can guide future implementation research [12] to, for example, examine how this program serves system needs and explore how the program can be improved through minor adjustments. Ultimately, understanding the processes and outcomes of the VA transgender e-consult program is critical for measuring its impact on enhancing care for transgender veterans. This project aimed to describe providers' program experiences through qualitative analysis, determine the reasons for nonuse of the program, and identify methods for improving program use through a quantitative survey.

Methods

Ethical Considerations

All qualitative activities were conducted as part of a quality-improvement initiative by the directors of VA's LGBT Health Program in the Office of Patient Care Services; therefore, no subject approval or informed consent to participate was required. The quantitative analyses were approved by the Institutional Review Board of the VA Pittsburgh Healthcare System.

Qualitative Methods

We received the LGBT Health Program's list of unique VA providers who used the VA transgender e-consult program, from the program's national launch in April 2015 through December 2016. We categorized the provider roster by urban and rural sites based on urban-rural definitions of the VA Office of Rural Health [13]. Subsequently, we randomly selected 32 users from urban sites and 34 users from rural sites. In January 2017, the selected providers received a recruitment email from the LGBT Health Program that explained the project goals and invited them to participate in a 20- to 30-minute semistructured phone interview to discuss their experiences with the transgender e-consult program. Approximately 1 week later, a project team member sent a follow-up email to the pertinent providers and reached out via direct phone calls. These efforts resulted in interviews with 3 urban and 6 rural providers. In May 2017, this process was repeated for providers who had not yet participated in the interviews, resulting in additional interviews with 1 urban and 5 rural providers (a total of 4 urban and 11 rural provider interviews).

With guidance from personnel who direct and administer the e-consult program, we developed a semistructured interview guide based on the most-germane needs identified. The initial set of guiding questions were as follows: How did providers learn about the availability of e-consultation? How useful was the e-consultation response? How did the treatment plan change as a result of the consultation? Were there provider-perceived changes in patient outcomes as a result of the consultation? What were the challenges or limitations of the e-consultation program? From the guiding questions, the project team developed and finalized an interview guide with open-ended questions and additional probes to direct in-depth inquiry. From January to May 2017, a trained interviewer conducted interviews, which were audio-recorded and transcribed for analysis.

Microsoft Word 2016 (Microsoft Corp., Redmond, WA) was used to transcribe the audio recordings verbatim to identify categories across interviews. The interviews were analyzed using conventional qualitative content analysis [14]. Using inductive analysis, the narrative text was reviewed using open coding. The coder read the interview transcripts multiple times to develop and refine categories in order to describe health care providers' experiences using the e-consult program. The lists of categories were sorted and grouped according to similar content for various aspects of providers' experiences in using the program under higher-order headings, with reduction of the data by content areas.

We performed quality control and bias mitigation in several ways. First, an experienced qualitative researcher reviewed the codebook prior to completion of the coding process. Second, the same researcher reviewed 20% (4 transcripts) of the samples of coded transcripts to ensure transparency and comprehensibility in the application of codes. The results were subsequently reviewed and discussed by three of the authors before their finalization.

Quantitative Methods

For the survey portion of our project, VA's administrative data were extracted from April 1, 2015, through April 1, 2017, for all inpatient and outpatient visits in which a transgender-related International Classification of Diseases (ICD)-9 or ICD-10 diagnosis code was noted for the visit (eg, gender-identity disorder, transsexualism, and personal history of sex reassignment). For each visit, the names of all unique providers who noted the diagnosis code were extracted to create a census of providers who encountered a transgender patient during the study period. The list of providers who had used the transgender e-consult program during the study period was cross-referenced with the total list of providers who encountered a transgender patient, and the providers who used the e-consult were removed. The remaining providers represented the eligible sample of providers who had encountered transgender patients during the study period, but did not utilize the e-consult program ($n=14,502$). Of the eligible providers, a random sample of 300 providers was selected, regardless of the provider type (eg, social worker or primary care physician). Email addresses of all providers were manually collected from the VA's global contact list, and addresses of 279 providers (of 300, 93%) were available, which comprised the analytical sample for this project. Providers received a recruitment email on November 13, 2017, which explained the purpose of the research project and provided a link to a brief 9-item survey on VA's internal network via REDCap (Vanderbilt University, Nashville, TN). Two weeks later (November 27), a second recruitment email was sent to all providers who had not completed the survey.

In the brief online survey, VA providers were asked to identify their discipline from social work, psychology, psychiatry, primary care, primary care nursing, endocrinology, speech therapy, nursing, and other. Providers indicating "other" were asked to elaborate on their answer. Before beginning the project with the VA, providers were asked if they ever received any training on providing health care to transgender patients; response options were yes, >1 hour; yes, >1 to <3 hours; yes, ≥ 3 hours; and no. Using the same response options, participants were asked if they ever received training when working for the VA. In addition, providers indicated the number of transgender patients they had cared for in the last 12 months. Finally, providers were asked, "Based on your own perception of the VA facility in which you do most of your work, how would you classify your VA facility?" The response options were urban, rural, or highly rural.

Three questions focused on the transgender e-consult program. First, providers indicated if they had ever used the program. For individuals who indicated "yes," they were thanked for their participation, and the survey ended. Individuals who indicated that they had not used the e-consult program received two additional questions. The first queried the providers about the issues that prevented them from using the program, and the response options included the following: I did not know about it, I do not know how to access it, I have not needed to use it yet, I am comfortable with my level of knowledge with transgender care, my patient(s) told me they did not want me to use it, I have a colleague I can call on for help, template takes too long to complete, and other. The second item queried

providers about the factors that would encourage them to use the program. Response options included the following: provision of more information about where to find e-consult in the chart, allotment of more time to complete an e-consult, encountering more transgender patients, VA-wide email announcements about the program, guidance on how to talk with patients about the program, and other. For both items, providers could choose more than one option, and persons indicating “other” were asked to elaborate on their answer.

Univariate statistics were used to describe the sample. Chi-square tests of independence and *t*-tests were used to examine categorical and mean differences, respectively. All quantitative analyses were conducted using Stata or SE, version 14.2 (StataCorp, College Station, TX). Values of $P < .05$ were considered statistically significant.

Results

Qualitative Evaluation

Participant Characteristics

Of the 15 interview participants included, most were primary care providers (12/15), practiced at VA for >10 years (7/15), categorized their current VA facility as rural (9/15), had no transgender-specific training (9/15), encountered 1 transgender patient in the last year (6/15), and reported that the program was 90%-100% useful for their specific needs (7/15) (Table 1). In addition, 27 of the total consultations were e-consults, and the majority (9/15) of the providers used the program at least once.

Fourteen VA health care providers participated in a telephone-based interview, of which 1 participant sent written responses via email because they could not schedule an interview. The longest interview lasted for 30 minutes 19 seconds, and the shortest was of 7 minutes 36 seconds; the mean interview time was 18 minutes 9 seconds. Six major areas were identified: how providers learned about the e-consult program, reasons for using the program, results of using the program, facilitators for using the program, barriers to using the program, and suggestions for improving the program. Table 2 includes exemplar quotes for each of the following categories identified during qualitative analyses.

Ways Providers Learned of the E-Consult Program

Providers commonly received information about the program through an email from the VA’s LGBT Health Program. Local providers or colleagues such as pharmacists, mental health providers (eg, psychologists), and gynecologists also educated interviewees before the e-consult as well as when another provider asked for assistance or advice through “word of mouth.” In addition, providers learned about the program by seeing “Transgender E-Consultation” as an option in the VA electronic health record (EHR). Some providers saw this option when entering an e-consult request in a different specialty area, whereas others saw it while going through orders. Other providers learned about the program through participation in

local VA LGBT-specific activities or educational lectures for providers about transgender care.

Reasons for Using the E-consult Program

Providers used the program to create individual patient-care plans, particularly providers with “extremely limited” transgender care experience. Some examples focused on gender-affirming medical care, including determining whether the patient was a good candidate for medical transition and the type of monitoring necessary during the transition process. Others sought recommendations about hormone therapies. Although there were “general medication questions,” most asked specific questions about appropriate medication doses and adjustments to optimize therapeutic outcomes and address safety concerns (eg, potential adverse effects). Furthermore, providers used the consult for education about sexual and reproductive health to help make care decisions such as those regarding standard preventative health-screening tests (eg, mammograms for transgender women) and gynecology for transgender men. Others desired education about available surgical treatment options (eg, breast augmentation and bilateral mastectomy). Moreover, providers sought guidance for communication with transgender patients on, for example, information about hormone therapy they should provide their patients and the responses they should expect. E-consults also provided a second opinion when providers were uncertain about a health care decision, especially in complex cases or cases involving high risk. The consults were found to be especially helpful when they confirmed what the provider already informed the patient or helped bolster provider decisions that were made against patient preferences.

Results of Using the E-consult Program

Providers sometimes modified care plans for their transgender patients after obtaining an e-consult response, including a formal diagnosis. Modifications included addition of recommended gender dysphoria evaluations, health screenings, medication monitoring, and meeting standard time frames for follow-up care. Most treatment plan changes were related to dosage of hormone therapy and addition of specialist referrals. Providers who used the program became more comfortable and confident in their assessments and in relaying testing and treatment decisions for transgender patients. Some providers felt confident in their judgment before the e-consult, but a second opinion confirmed what they discussed and decided with the patient. E-consults helped improve patient-provider relationships by placing those interactions and care provision in line with the gender (ie, body, identity, and expression) of their patients. In addition, the e-consults provided patients added motivation to follow recommendations because experienced providers were involved. Moreover, communication during clinic interactions improved because the e-consult request requires providers to answer a number of questions about the patient, which encouraged them to collect additional relevant information from the patient during their visit. Discussing recommendations with their patients allowed them to engage in informed, shared decision making about their care.

Table 1. Characteristics of the 15 Department of Veterans Affairs (VA) providers from qualitative interviews.

Characteristics	Providers, n (%)
Type of VA health care provider	
Primary care	12 (80)
Mental health	3 (20)
Transgender-specific training (in VA and nonVA settings)	
No	9 (60)
Yes	6 (40)
Years in practice at VA	
0-3 years	3 (20)
>3-5 years	2 (13)
>5-7 years	1 (7)
>7-10 years	2 (13)
>10 years	7 (47)
Number of transgender patients in the last year	
1	6 (40)
2	3 (20)
3	3 (20)
4	2 (13)
>5	1 (7)
Number of e-consults performed	
1	9 (60)
2	4 (27)
3	1 (7)
4	0 (0)
5	0 (0)
6	0 (0)
7	1 (7)
Usefulness of the e-consult program	
>90%-100%	7 (47)
>80%-90%	4 (27)
>70%-80%	1 (7)
>60%-70%	1 (7)
>50%-60%	0 (0)
≤50%	1 (7)
No answer	1 (7)
Location of facility (participant self-report)	
Rural	9 (60)
Urban	3 (20)
Suburban or mixed	2 (13)
No answer or not asked	1 (7)
Location of facility (defined by the Office of Rural Health)	
Rural	11 (73)
Urban	4 (27)

Table 2. Qualitative findings among providers using the Department of Veterans Affairs (VA) transgender e-consult program.

Categories	Quotation examples
Topic 1: Ways providers learned of the program	
Email notification	“There was an email with a flyer.”
Informed by other staff members	“My pharmacist or my mental health provider at the clinic showed me how to access that consult.”
Via the EHR ^a system	“Through the VA Consult Program, ...through the orders, you can go to the e-consult request and order sections in the records and there’s ‘Transgender E-Consults’ you can click on.”
Participation in VA LGBT ^b activities and educational trainings	“Through one of those lectures that we got about transgender care.”
Topic 2: Reasons for using the program	
Establish a care plan	“My experience with the gender reassignment is extremely limited.”
Recommendations about hormone-replacement therapies	“I was having too much testosterone and I needed to...see if I was doing something not quite right. They gave me some tips like, ‘Yes, we want this level for the testosterone in order for...transition’.”
Education about sexual and reproductive health care	“I wasn’t sure how to screen her because I didn’t know...how often to get her a mammography.”
Education about surgical treatment options	“What options we might have available for surgical interventions.”
Guidance regarding patient-provider communication	“I typically will ask about what types of responses should I expect...What should I tell patients about what type of breast size increase they may expect.”
Second opinion	“This fellow is a senior and he’s got multiple medical problems and he’s on anticoagulants, and lots of cardiac issues, and hormones would be a really bad idea. (I used the consult to) just evaluate him and see if he’s a candidate.”
Topic 3: Results of using the program	
Care-plan modifications	“It was helpful for outlining starting medicines and then following the intervals for following those with laboratories.”
Increased provider comfort with and confidence in transgender care provision	“They actually had a template to walk someone through it. And it was exceptional. Because, while I have worked with this population before, I really, at that point, wasn’t comfortable starting from scratch in the assessment.”
Improved patient-provider relationships	“I thought (the e-consult) did ask questions that kind of allowed me to ask the patient more questions. So, it kind of delved a bit, so that if you are new to it...”
Continued use of the e-consult program	“I’ve only offered it to the one patient, but on multiple occasions.”
Topic 4: Facilitators for using the program	
Responses contain understandable and informative recommendations	“I do exactly what they say, and if it doesn’t work they say ‘This is the next step you should take.’ They’ll say, ‘Go ahead and give us a call back.’”
Accessibility and ease of program use by providers through the EHR	“It was just on our primary (EHR) page, so it was really easy to find... When we put in for consults, it’s on the very main screen.”
Quick response to e-consult request	“They answered my inquiry right away.”
Provider knowledge of the program’s existence	“Knowing it was out there.”
Only available transgender-specific resource for providers at rural sites	“I don’t have much of a choice. The only resource we have is to use the e-consult.”
Topic 5: Barriers to using the program	
Time-consuming process for the provider to submit an e-consult and read through recommendations	“They pull in parts of notes from the mental health evaluation, from my evaluation, from all sorts of things. So, the note is not clear when it comes back... They probably have a dozen pages where they (explain) hormone levels, when they should be checked... It’s meant to prevent re-consult, which is frustrating. It’s very cumbersome...all that stuff just automatically gets tacked onto the bottom of the consult.”
Previous e-consult response did not answer the question asked	“One of the responses I received in the e-consult was that I should become an advocate and get more programs. And, honestly, that answer left me feeling demoralized. It was just unrealistic given my workload.”
Communication-related problems with e-consult	“It’s not really clear from the consult, they want you to submit another e-consult...sometimes you may, six months later, have a question...that may require the submission of a new e-consult (rather than just adding a comment to the existing consult).”

Categories	Quotation examples
Impractical e-consult recommendations	“The recommendation (to) see the patient back in a couple months and the follow-up intervals they recommended, to be honest I can’t accommodate.”
Provider misunderstanding of e-consult purpose	“I just don’t know why they didn’t take the case and take over.”
Lack of provider understanding of transgender patients and their specific health care needs	“I had been kind of thinking of her as a man in my head... I had been seeing her for about a year before I had even thought about this... I think it was just a lack of awareness.”
Topic 6: Suggestions to improve the program	
Adding concise and sectioned e-consult responses	“Appropriate hormone levels, intervals for checking labs, any changes in preventative care, if those were in separate sections. Because now...it’s just kind of merged into one on-going paragraph.”
Expanding provider and patient awareness about the program	“Perhaps market it more...I’m pretty sure not a lot of people know it exists.”
Designating a contact for follow-up after the e-consult	“To have a contact person who I could have more of a dialogue with. There were multiple people contributing to the e-consult, but no one contact person that I could ask logistical questions.”
Increasing provider education about transgender veterans and transgender-specific health care	“Maybe a little more education on the consult...the consult education was provided after they...look at things and then they were like, ‘Oh, well, this was what you needed to do.’ But if, in the consult itself, you list all of the pre-requisites that are necessary (for transitioning).”
Clarification for providers on e-consult operational processes and timelines	“Some idea of the turnaround process would have been helpful.”

^aEHR: electronic health record.

^bLGBT: lesbian, gay, bisexual, and transgender.

Facilitators for Using the E-consult Program

E-consult responses contained “timely and up-to-date-information” that was useful and easily understood because it did not contain jargon. They appreciated the fact that the e-consult was reviewed by an expert multidisciplinary team in transgender care and saw them as “coaches.” The recommendations reportedly provided a clear “road map” for where to start and how to proceed with care. The program was said to be “easy to access” and use through the VA EHR system, which also made it “easier for continuity of care” because information such as test and treatment results can be added, accessed together, reviewed, and incorporated into a patient’s care plan.

Providers who used e-consults received responses in a “timely manner,” ranging from 2 days to 2 weeks, depending on the question, which was much shorter than the time for other types of consultations they requested in the past. The program made providers aware of other existing health care services and resources (eg, sexual and reproductive health care services). Some rural providers said they used the e-consult program because they believed it was their only available resource for patient-specific transgender health. In particular, providers at rural sites said there were no on-site providers or resources they could consult about care for transgender patients.

Barriers to Using the E-consult Program

The providers reported that entering the required information to submit an e-consult and reading the responses could be time consuming. For example, providers had to enter extensive patient information, some of which was not directly relevant to their question (eg, “Mental health things like suicide or homicide, interpersonal violence”). Some providers thought the template contained repetitive questions. Others said the

responses were too long (eg, “If you printed it off, it would probably be like 20 pages long”) and not user friendly, including templated information and irrelevant information that they would not use. One provider was frustrated when the response did not address her question and suggested unrealistic and broad changes in her practice.

Communication-related issues were noted with receipt of the response and follow-up questions, some of which related to the technology central to e-consults. For example, in a situation with multiple e-consults for different questions about the same patient over time, it was unclear whether providers should add a new e-consult or add comments to their previous e-consult. Providers were also unable to follow recommendations because they found them impractical or unfeasible (eg, too busy due to their existing workload). Some providers at rural facilities could not follow recommendations because they lacked the necessary local resources, and the e-consult did not help them access the required resources. A few providers misunderstood the e-consult purpose, which led them to misuse the e-consult and question why their request was not met. For example, some providers thought a referral request would result in other providers taking over their patient’s care. Some providers did not use e-consult due to a lack of understanding about transgender patients, including dimensions of gender (ie, body, identity, and expression) and transgender-specific health care needs.

Suggestions for Improving the E-consult Program

Providers suggested streamlining responses into sections to make them easier to read and understand. Redundant or templated sections could be moved to a common share point on the VA intranet. Additionally, providers suggested increasing provider and patient awareness of the program; for example, “Getting information out that, if the veteran has concerns, they

can go to their primary care clinics and get the process started, so they don't have to go outside the VA." Providers also suggested designating a point-person after an e-consult is completed, so that providers can follow-up and discuss the issue, and not just receive a one-sided, one-time response. Further, they suggested provision of additional provider education about transgender-specific health care needs through the e-consult and the VA Specialty Care Access Network-Extension for

Community Healthcare Outcomes program. Creation of new programs would help educate providers about transgender veterans, including dimensions of gender and transgender-specific health care. Providers suggested that additional information and clarification be given to providers about the e-consult program's operational processes and timeline for e-consults, including the estimated time required to enter an e-consult request and average wait time for a response.

Table 3. Comparison between urban and rural providers in 53 provider surveys.

Parameter	Urban (n=35)	Rural (n=18)	P value
Discipline, n (%)			
Social work	6 (17.6)	2 (11.1)	.82
Psychology	7 (20.6)	3 (16.7)	— ^a
Psychiatry	4 (11.8)	3 (16.7)	—
Primary care	5 (14.7)	3 (16.7)	—
Nursing	5 (14.7)	4 (22.2)	—
Other	7(20.6)	3 (16.7)	—
Training in transgender health prior to working with the VA^b, n (%)			
Yes, 1 hour	5 (14.7)	1 (5.6)	.72
Yes, >1 to <3 hours	8 (23.5)	3 (16.7)	—
Yes, >3 hours	4 (11.8)	3 (16.7)	—
None	17 (50.0)	11 (61.1)	—
Training in transgender health while working with the VA, n (%)			
Yes	19 (57.6)	7 (38.9)	.20
No	14 (42.4)	11 (61.1)	—
Prior use of transgender e-consult, n (%)			
Yes	2 (5.9)	3 (16.7)	.33
No	32 (94.1)	15 (83.3)	—
Reasons for nonuse of the e-consult, n (%)			
Did not know about it	17 (50.0)	12 (66.7)	.25
Do not know how to access it	7 (20.6)	5 (27.8)	.56
Have not needed to use it yet	12 (35.3)	5 (27.8)	.58
Comfortable with knowledge of transgender care	6 (17.6)	1 (5.6)	.40
Patient(s) told me they did not want me to use it	0	0	—
I have a colleague I can call on for help	11 (32.3)	2 (11.1)	.18
Template takes too long to complete	0	0	—
Suggestions to encourage use of the e-consult, n (%)			
More information about where to find e-consult in the chart	18 (52.9)	11 (61.1)	.57
Have more time to complete an e-consult	6 (17.6)	4 (22.2)	.72
Seeing more transgender patients	10 (29.4)	5 (27.8)	.90
VA-wide email announcements about the program	14 (41.2)	9 (50.0)	.54
Guidance on how to talk with patients about the program	17 (50.0)	8 (44.4)	.70
Number of years working in VA, mean (SD)	7.0 (0.8)	6.8 (1.4)	.45

^aNot applicable.

^bVA: Department of Veterans Affairs

Quantitative Evaluation

Overall Analysis

Comparison of the three demographic characteristics between the census of providers who had encountered a transgender patient during the study period and the random sample of survey participants showed no significant differences in facilities ($P=.36$), number of visits ($P=.48$), and number of patients ($P=.31$) (data not shown). Of the 279 eligible providers, 53 responded to the survey (19% response rate). Providers were evenly distributed across disciplines (Table 3).

Among primary care practitioners, 5 were physicians, 3 were nurse practitioners, and 10 were “other” (ie, nurse practitioner in gynecology, art therapist, blind rehabilitation, clinical pharmacy specialist, physician in physical medicine and rehabilitation, nonspecific primary care provider, certified nutrition specialist, psychiatric nurse practitioner, registered dietitian nutritionist, vascular surgeon, or inpatient physician). On an average, the practitioners worked in the VA for 6.9 years (SD 5.3; range, ≤ 1 -25 years). A total of 35 respondents (66%) viewed their VA facility as urban; 15 (28%), as rural; and 3 (6%), as highly rural.

Over half (53%) of the sample had no training in providing transgender health care before working in the VA. Among those who had prior training, 11% had 1 hour, 21% had >1 to <3

hours, and 15% had >3 hours of training. In addition, half of the sample (50%) indicated that while they were working for the VA, they received some form of training on providing care for transgender patients.

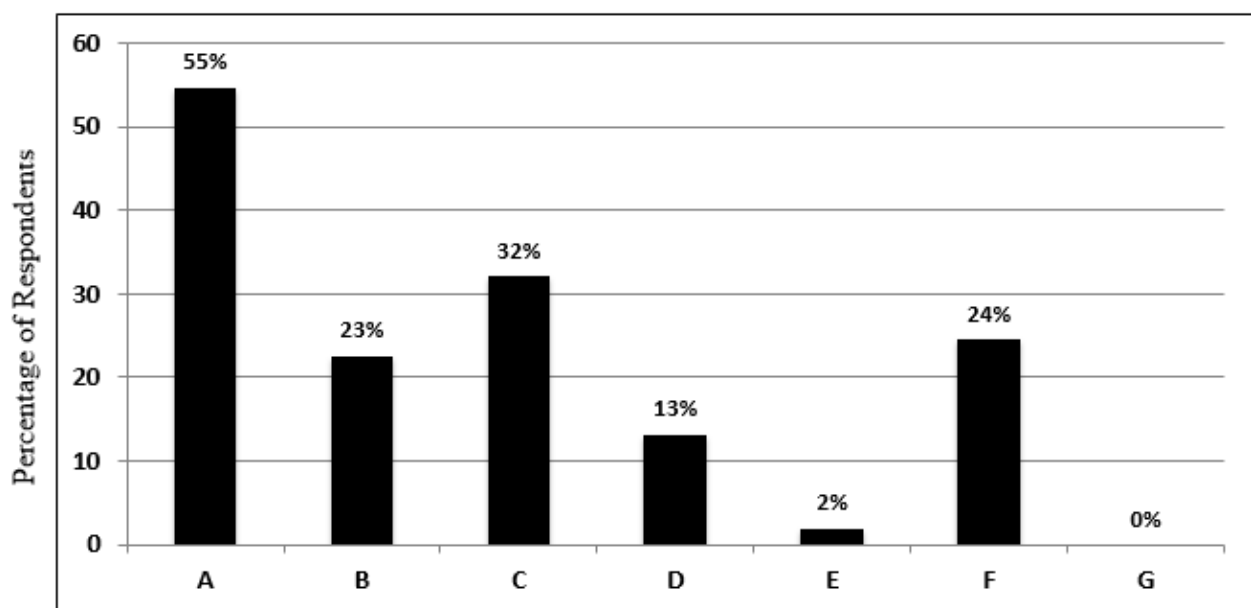
Nearly one-third (32%) of the providers encountered only 1 transgender patient in the previous 12 months, and 5 providers (9%) did not encounter any transgender patients in the last 12 months. The majority of the providers (87%) had encountered 1-9 transgender patients over the last 12 months, and 2 providers encountered >10 transgender patients.

The most-common reasons for nonuse of the e-consult were as follows: no knowledge about it, no need for it, and receipt of help from a colleague (Figure 1). The most-common suggestions to improve the program use were provision of more information about how to find and access e-consult through the EHR, VA-wide email announcements about the program, and guidance on how to talk with patients about the program (Figure 2).

Urban and Rural Differences

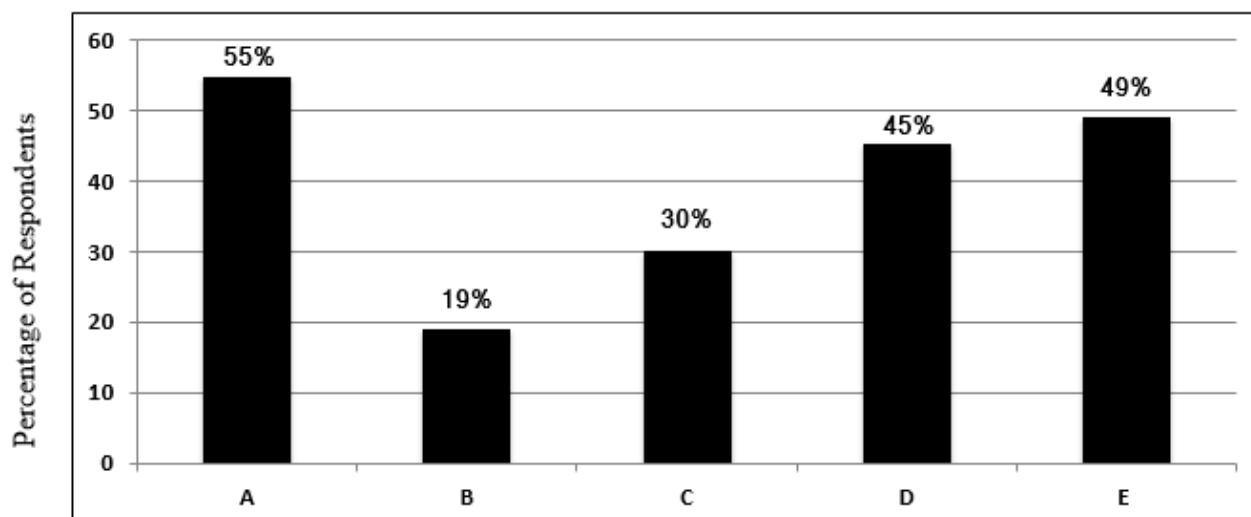
Post hoc exploratory analyses were conducted on the basis of the self-reported locale of the providers. Rural and highly rural respondents were combined in one category because of the small number of providers in highly rural facilities. We found no overall differences between urban and rural providers (Table 3).

Figure 1. Reasons for not using e-consultation from 53 provider surveys.



A=did not know about program; B=did not know how to access; C=have not needed to use it yet; D=comfortable with my knowledge of transgender care; E=patient did not want me to use it; F=call on a colleague; G=template takes too long to complete

Figure 2. How to improve use of e-consultation from 53 provider surveys. VA: Department of Veterans Affairs.



A=more information about where to find e-consult in the chart; B=have more time to complete an e-consult; C=seeing more transgender patients; D=VA-wide email announcements about the program; E=Guidance on how to talk with patients about the program

Discussion

Principal Findings

Increase of provider knowledge about transgender patients' needs can reduce misconceptions and stigma and facilitate high-quality care provision [1,15]. The VA has improved transgender health care by becoming the first and largest US health care system to design and implement a national e-consult program for transgender health care to "respond to specific inquiries on a case-by-case basis" [10].

Because formal transgender health care is relatively new to the VA, it is crucial to understand health care providers' experiences with a program designed to enhance their practice. Lessons learned can guide improvement and inform other health care organizations that may want to develop or replicate such a program. Some large health care organizations have implemented e-consult programs for several system needs, mainly those facilitating low-cost connections with specialists and continuity of care within the system [16-18]. The implementation of e-consults as a relatively inexpensive, structural service that consolidates clinical expertise into an accessible resource can be an asset for vulnerable populations and providers who serve them. Our study revealed the abovementioned assets among providers caring for patients who are transgender (eg, increased provider comfort with and confidence in transgender care provision) and elucidated the importance of the program in ensuring flexibility in response to provider's dynamic needs (eg, structured responses to reduce response time).

Because the VA operates the single-largest health care system in the United States, we were able to specifically examine the experiences of providers at rural sites. Qualitative data showed that rural providers who have used e-consult faced unique challenges in providing quality care to transgender veterans. In particular, they did not have on-site providers with knowledge

of and experience in transgender-specific care, who they could consult about the veteran. At times, they were unable to follow the e-consult recommendations they received for a specific patient, because the recommendations were not feasible at their location. Thus, the e-consult program did not provide local resources needed to follow providers' advice. These themes highlight a conundrum that required more directed inquiry, such as exploring how providers in rural areas can deliver the same quality of care as their urban peers, consolidating e-consult responses to rural providers in order to develop tailored guidelines for providing transgender care in underresourced areas, or convening system-focused consensus meetings to bring together providers and administrators from both urban and rural settings to discuss creative quality-improvement strategies.

Although our quantitative analyses showed no statistically significant differentiation between urban and rural providers, the results suggested that rural respondents may be less likely than their urban peers to indicate that they have colleagues who can be informally called for help, reiterating that rural areas generally have limitations to health care infrastructure and availability of providers [19-21], which could have negative effects for care of unique minority populations such as transgender individuals [22-24]. Further research is needed to develop a better understanding of transgender health from the inputs of both providers and patients, specifically in rural settings.

Our results are consistent with existing research findings that highlight ongoing barriers to provision of quality transgender-specific care in VA and nonVA settings [25-28]. Although continually revised standards of care for transgender patients have been available through the World Professional Association for Transgender Health for over 2 decades [29], the majority of health care professionals do not receive training in transgender health [1]. For instance, in a survey of 132 deans representing medical schools in the United States and Canada, one-third reported 0 hours of LGBT-related content delivered during

clinical training years and only 30% reported inclusion of gender transition in the required curriculum [30]. In a sample of nearly 250 third- and fourth-year medical students, only 56% chose the correct definition of the term “transgender” [31]. Unsurprisingly, some of the frequently cited barriers to accessing care among transgender patients are the lack of knowledgeable providers (regarding body, identity, expression, etc, for transgender patients and transgender-specific care) [15,32,33] or the need for the patient to educate the provider [4], which may be magnified for patients in rural and remote locations [22,34]. Although medical training and curriculum need to include transgender topics [35], health systems face a concomitant, immediate need to equip current providers with necessary resources in order to provide transgender-specific health care for transgender patients.

Limitations

We noted several limitations to our study. The qualitative analyses relied on self-reported data from interviews with providers who used the VA transgender e-consult program, which may not replicate direct observational data. Although our relatively small and homogeneous sample was suitable for qualitative analysis, we may not have captured the full range of provider experiences with the VA transgender e-consult program. Notably, the majority of the sample had only used the e-consult program once and one person had used it 7 times; perceptions of the program could vary based on familiarity. In addition, due to the qualitative nature of the study, our results may not be representative or generalizable; this limitation should be considered when interpreting the results before applying our conclusions to other providers and settings.

For the quantitative survey, we identified providers who encountered transgender patients (ie, the use of ICD-9 and ICD-10 diagnosis codes), which is a proxy method of identifying transgender patients; therefore, the sampling may have resulted in a conservative estimate of providers. The quantitative response rate was low, limiting the statistical power of analyses and generalizability of findings. Moreover, because we did not have data on the disciplines for the 14,502 providers in the pool that was sampled, we could not ascertain if providers from certain disciplines were more inclined to respond to the survey, which may have introduced response bias. Future studies with more-frequent and varied follow-up strategies to increase the

response rate could alleviate issues of generalizability and response bias. Furthermore, as a pilot study, the survey was limited in scope and did not collect additional information about providers (eg, sex, race or ethnicity, and granular data regarding specialty training).

Conclusions

This mixed-methods study offers insight into provider experiences with the VA transgender e-consult program and conveys lessons of implementation that can guide other health care organizations to create similar programs (eg, better and more-frequent advertising of the program, infographics or examples about how to initiate an e-consult, or structured advice for providers to talk with patients about e-consults). Our results suggest that the program is useful for providers, but there are several barriers to its use among providers. Additionally, rural providers face unique challenges in care provision for transgender patients, including the lack of access to necessary resources.

Future studies with a larger, more-diverse sample of urban and rural health care providers as well transgender veterans who were involved in the e-consult program (patients must consent for their providers to submit an e-consult) are warranted. The present study included only the providers who asked for e-consults and not the providers who answered the e-consults. Further research, probably via a dyadic approach, could better investigate the entire course of e-consult use. Additionally, a global study of VA e-consults would allow comparisons across consult content (eg, diabetes, dermatology-related, and transgender care) to discern challenges that may be universal to the e-consult format in contrast to challenges that may be unique based on the health care need. Examining transgender e-consults from the patient’s perspective would allow the health system and its providers to learn how consumer experience (eg, patient-provider interactions, patient outcomes, or satisfaction) may be influenced by the use of e-consults [36].

In summary, an e-consult is a relatively low-cost system improvement that health care organizations can implement to support providers in caring for transgender patients. As examples of best-practice suggestions for transgender health accumulate [37-39], feasible, in-house resources will be key for their replication and adaptation across health care organizations.

Acknowledgments

This work was supported by a pilot research award from the Department of Veterans Affairs (VA) Center for Health Equity Research and Promotion Competitive Pilot Program (PI: JRB; XIP-72-006) and the VA Health Services Research and Development Career Development Award (PI: JRB; CDA-14-408). The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the VA or the United States Government.

Authors' Contributions

JRB, KLR, AJG, AM, JCS, MRK, and SMM contributed to the project concept and design. KLH conducted the interviews. KLH and DK coded the interviews, and KLH and KLR completed the qualitative analysis. JBR completed all quantitative analyses. All authors contributed substantially to the manuscript preparation and review.

Conflicts of Interest

None declared.

References

1. Institute of Medicine. The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding. Washington, DC: Institute of Medicine; 2011:13.
2. Stotzer RL. Violence against transgender people: A review of United States data. *Aggression and Violent Behavior* 2009 May;14(3):170-179. [doi: [10.1016/j.avb.2009.01.006](https://doi.org/10.1016/j.avb.2009.01.006)]
3. Brown GR, Jones KT. Mental Health and Medical Health Disparities in 5135 Transgender Veterans Receiving Healthcare in the Veterans Health Administration: A Case–Control Study. *LGBT Health* 2016 Apr;3(2):122-131. [doi: [10.1089/lgbt.2015.0058](https://doi.org/10.1089/lgbt.2015.0058)]
4. Grant J, Mottet L, Tanis J, Herman J, Harrison J, Keisling M. National Transgender Discrimination Survey Report on Health Care. Washington, D. C.: National Center for Transgender Equality and the the National Gay and Lesbian Task Force; 2010. URL: <https://cancer-network.org/wp-content/uploads/2017/02/National-Transgender-Discrimination-Survey-Report-on-health-and-health-care.pdf> [WebCite Cache ID 72kLetqZc]
5. Legato MJ, editor. Standards of Care for the Health of Transsexual, Transgender, and Gender-Nonconforming People: An Introduction. In: *Principles of Gender-Specific Medicine: Gender in the Genomic Era*. Cambridge, MA: Academic Press; 2017:69-75.
6. Lutwak N, Byne W, Erickson-Schroth L, Keig Z, Shipherd J, Mattocks K, et al. Transgender veterans are inadequately understood by health care providers. *Mil Med* 2014;179(5):483-485. [doi: [10.7205/MILMED-D-14-0000](https://doi.org/10.7205/MILMED-D-14-0000)]
7. Department of Veterans Affairs, Veterans Health Administration. VA Health Care. 2018. VHA strategic plan FY 2013-2018 URL: https://www.va.gov/health/docs/VHA_STRATEGIC_PLAN_FY2013-2018.pdf [accessed 2018-09-27] [WebCite Cache ID 72kMSNQ2c]
8. Blosnich JR, Brown GR, Shipherd, PhD JC, Kauth M, Piegari RI, Bossarte RM. Prevalence of Gender Identity Disorder and Suicide Risk Among Transgender Veterans Utilizing Veterans Health Administration Care. *Am J Public Health* 2013 Oct;103(10):e27-e32. [doi: [10.2105/AJPH.2013.301507](https://doi.org/10.2105/AJPH.2013.301507)]
9. Kauth MR, Shipherd JC, Lindsay J, Blosnich JR, Brown GR, Jones KT. Access to Care for Transgender Veterans in the Veterans Health Administration: 2006–2013. *Am J Public Health* 2014 Sep;104(S4):S532-S534. [doi: [10.2105/AJPH.2014.302086](https://doi.org/10.2105/AJPH.2014.302086)]
10. Shipherd JC, Kauth MR, Matza A. Nationwide Interdisciplinary E-Consultation on Transgender Care in the Veterans Health Administration. *Telemedicine and e-Health* 2016 Dec;22(12):1008-1012. [doi: [10.1089/tmj.2016.0013](https://doi.org/10.1089/tmj.2016.0013)]
11. Vimalananda VG, Gupte G, Seraj SM, Orlander J, Berlowitz D, Fincke BG, et al. Electronic consultations (e-consults) to improve access to specialty care: A systematic review and narrative synthesis. *J Telemed Telecare* 2015 May 20;21(6):323-330. [doi: [10.1177/1357633X15582108](https://doi.org/10.1177/1357633X15582108)]
12. Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, et al. The Role of Formative Evaluation in Implementation Research and the QUERI Experience. *J Gen Intern Med* 2006 Feb;21(S2):S1-S8. [doi: [10.1111/j.1525-1497.2006.00355.x](https://doi.org/10.1111/j.1525-1497.2006.00355.x)] [Medline: [16637954](https://pubmed.ncbi.nlm.nih.gov/16637954/)]
13. West AN, Lee RE, Shambaugh-Miller MD, Bair BD, Mueller KJ, Lilly RS, et al. Defining “rural” for veterans' health care planning. *J Rural Health* 2010;26(4):301-309. [doi: [10.1111/j.1748-0361.2010.00298.x](https://doi.org/10.1111/j.1748-0361.2010.00298.x)] [Medline: [21029164](https://pubmed.ncbi.nlm.nih.gov/21029164/)]
14. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008 Apr;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
15. Safer JD, Coleman E, Feldman J, Garofalo R, Hembree W, Radix A, et al. Barriers to healthcare for transgender individuals. *Curr Opin Endocrinol Diabetes Obes* 2016 Apr;23(2):168-171 [FREE Full text] [doi: [10.1097/MED.0000000000000227](https://doi.org/10.1097/MED.0000000000000227)] [Medline: [26910276](https://pubmed.ncbi.nlm.nih.gov/26910276/)]
16. Liddy C, Drosinis P, Keely E. Electronic consultation systems: worldwide prevalence and their impact on patient care-a systematic review. *Fam Pract* 2016 Dec;33(3):274-285. [doi: [10.1093/fampra/cmw024](https://doi.org/10.1093/fampra/cmw024)] [Medline: [27075028](https://pubmed.ncbi.nlm.nih.gov/27075028/)]
17. Nijland N, van Gemert-Pijnen JEWC, Boer H, Steehouder MF, Seydel ER. Increasing the use of e-consultation in primary care: results of an online survey among non-users of e-consultation. *Int J Med Inform* 2009 Oct;78(10):688-703. [doi: [10.1016/j.ijmedinf.2009.06.002](https://doi.org/10.1016/j.ijmedinf.2009.06.002)] [Medline: [19625210](https://pubmed.ncbi.nlm.nih.gov/19625210/)]
18. Tuot DS, Leeds K, Murphy EJ, Sarkar U, Lyles CR, Mekonnen T, et al. Facilitators and barriers to implementing electronic referral and/or consultation systems: a qualitative study of 16 health organizations. *BMC Health Serv Res* 2015 Dec 19;15:568 [FREE Full text] [doi: [10.1186/s12913-015-1233-1](https://doi.org/10.1186/s12913-015-1233-1)] [Medline: [26687507](https://pubmed.ncbi.nlm.nih.gov/26687507/)]
19. Johnson ME, Brems C, Warner TD, Roberts LW. Rural-urban health care provider disparities in Alaska and New Mexico. *Adm Policy Ment Health* 2006 Jul;33(4):504-507. [doi: [10.1007/s10488-005-0001-7](https://doi.org/10.1007/s10488-005-0001-7)] [Medline: [16220242](https://pubmed.ncbi.nlm.nih.gov/16220242/)]
20. Laditka JN, Laditka SB, Probst JC. Health care access in rural areas: evidence that hospitalization for ambulatory care-sensitive conditions in the United States may increase with the level of rurality. *Health Place* 2009 Sep;15(3):731-740. [doi: [10.1016/j.healthplace.2008.12.007](https://doi.org/10.1016/j.healthplace.2008.12.007)] [Medline: [19211295](https://pubmed.ncbi.nlm.nih.gov/19211295/)]

21. Douthit N, Kiv S, Dwolatzky T, Biswas S. Exposing some important barriers to health care access in the rural USA. *Public Health* 2015 Jun;129(6):611-620. [doi: [10.1016/j.puhe.2015.04.001](https://doi.org/10.1016/j.puhe.2015.04.001)] [Medline: [26025176](https://pubmed.ncbi.nlm.nih.gov/26025176/)]
22. Knutson D, Koch J, Arthur T, Mitchell T, Martyr M. Trans broken arm: Health care stories from transgender people in rural areas. *J Res Woman Gender* 2017;7:30-46.
23. Horvath KJ, Iantaffi A, Swinburne-Romine R, Bockting W. A comparison of mental health, substance use, and sexual risk behaviors between rural and non-rural transgender persons. *J Homosex* 2014 Jun;61(8):1117-1130 [FREE Full text] [doi: [10.1080/00918369.2014.872502](https://doi.org/10.1080/00918369.2014.872502)] [Medline: [24380580](https://pubmed.ncbi.nlm.nih.gov/24380580/)]
24. Bukowski LA, Blosnich J, Shipherd JC, Kauth MR, Brown GR, Gordon AJ. Exploring Rural Disparities in Medical Diagnoses Among Veterans With Transgender-related Diagnoses Utilizing Veterans Health Administration Care. *Med Care* 2017 Sep;55 Suppl 9 Suppl 2:S97-S103. [doi: [10.1097/MLR.0000000000000745](https://doi.org/10.1097/MLR.0000000000000745)] [Medline: [28806372](https://pubmed.ncbi.nlm.nih.gov/28806372/)]
25. Stroumsa D. The state of transgender health care: policy, law, and medical frameworks. *Am J Public Health* 2014 Mar;104(3):e31-e38. [doi: [10.2105/AJPH.2013.301789](https://doi.org/10.2105/AJPH.2013.301789)] [Medline: [24432926](https://pubmed.ncbi.nlm.nih.gov/24432926/)]
26. Vance SR, Halpern-Felsher BL, Rosenthal SM. Health care providers' comfort with and barriers to care of transgender youth. *J Adolesc Health* 2015 Feb;56(2):251-253. [doi: [10.1016/j.jadohealth.2014.11.002](https://doi.org/10.1016/j.jadohealth.2014.11.002)] [Medline: [25620310](https://pubmed.ncbi.nlm.nih.gov/25620310/)]
27. Kosenko K, Rintamaki L, Raney S, Maness K. Transgender patient perceptions of stigma in health care contexts. *Med Care* 2013 Sep;51(9):819-822. [doi: [10.1097/MLR.0b013e31829fa90d](https://doi.org/10.1097/MLR.0b013e31829fa90d)] [Medline: [23929399](https://pubmed.ncbi.nlm.nih.gov/23929399/)]
28. Xavier J, Bradford J, Hendricks M, Safford L, McKee R, Martin E, et al. Transgender Health Care Access in Virginia: A Qualitative Study. *International Journal of Transgenderism* 2013 Jan;14(1):3-17. [doi: [10.1080/15532739.2013.689513](https://doi.org/10.1080/15532739.2013.689513)]
29. Coleman E, Bockting W, Botzer M, Cohen-Kettenis P, DeCuypere G, Feldman J, et al. Standards of Care for the Health of Transsexual, Transgender, and Gender-Nonconforming People, Version 7. *International Journal of Transgenderism* 2012 Aug;13(4):165-232. [doi: [10.1080/15532739.2011.700873](https://doi.org/10.1080/15532739.2011.700873)]
30. Obedin-Maliver J, Goldsmith ES, Stewart L, White W, Tran E, Brenman S, et al. Lesbian, gay, bisexual, and transgender-related content in undergraduate medical education. *JAMA* 2011 Sep 07;306(9):971-977. [doi: [10.1001/jama.2011.1255](https://doi.org/10.1001/jama.2011.1255)] [Medline: [21900137](https://pubmed.ncbi.nlm.nih.gov/21900137/)]
31. Sanchez NF, Rabatin J, Sanchez JP, Hubbard S, Kalet A. Medical students' ability to care for lesbian, gay, bisexual, and transgendered patients. *Fam Med* 2006 Jan;38(1):21-27 [FREE Full text] [Medline: [16378255](https://pubmed.ncbi.nlm.nih.gov/16378255/)]
32. Rosentel K, Hill BJ, Lu C, Barnett JT. Transgender Veterans and the Veterans Health Administration: Exploring the Experiences of Transgender Veterans in the Veterans Affairs Healthcare System. *Transgend Health* 2016 Jan;1(1):108-116 [FREE Full text] [doi: [10.1089/trgh.2016.0006](https://doi.org/10.1089/trgh.2016.0006)] [Medline: [29159302](https://pubmed.ncbi.nlm.nih.gov/29159302/)]
33. Sanchez NF, Sanchez JP, Danoff A. Health care utilization, barriers to care, and hormone usage among male-to-female transgender persons in New York City. *Am J Public Health* 2009 Apr;99(4):713-719. [doi: [10.2105/AJPH.2007.132035](https://doi.org/10.2105/AJPH.2007.132035)] [Medline: [19150911](https://pubmed.ncbi.nlm.nih.gov/19150911/)]
34. Kauth MR, Barrera TL, Denton FN, Latini DM. Health Differences Among Lesbian, Gay, and Transgender Veterans by Rural/Small Town and Suburban/Urban Setting. *LGBT Health* 2017 Jun;4(3):194-201. [doi: [10.1089/lgbt.2016.0213](https://doi.org/10.1089/lgbt.2016.0213)] [Medline: [28430020](https://pubmed.ncbi.nlm.nih.gov/28430020/)]
35. Shipherd JC. Defining Competence when Working with Sexual and Gender Minority Populations: Training Models for Professional Development. *Clin Psychol Sci Pract* 2015 Jun 18;22(2):101-104. [doi: [10.1111/cpsp.12100](https://doi.org/10.1111/cpsp.12100)]
36. Lehavot K, Katon JG, Simpson TL, Shipherd JC. Transgender Veterans' Satisfaction With Care and Unmet Health Needs. *Med Care* 2017 Sep;55 Suppl 9 Suppl 2:S90-S96. [doi: [10.1097/MLR.0000000000000723](https://doi.org/10.1097/MLR.0000000000000723)] [Medline: [28806371](https://pubmed.ncbi.nlm.nih.gov/28806371/)]
37. Reisner SL, Bradford J, Hopwood R, Gonzalez A, Makadon H, Todisco D, et al. Comprehensive transgender healthcare: the gender affirming clinical and public health model of Fenway Health. *J Urban Health* 2015 Jun;92(3):584-592 [FREE Full text] [doi: [10.1007/s11524-015-9947-2](https://doi.org/10.1007/s11524-015-9947-2)] [Medline: [25779756](https://pubmed.ncbi.nlm.nih.gov/25779756/)]
38. Wylie K, Knudson G, Khan SI, Bonierbale M, Watanyusakul S, Baral S. Serving transgender people: clinical care considerations and service delivery models in transgender health. *The Lancet* 2016 Jul;388(10042):401-411. [doi: [10.1016/S0140-6736\(16\)00682-6](https://doi.org/10.1016/S0140-6736(16)00682-6)]
39. Deutsch MB, Radix A, Reisner S. What's in a Guideline? Developing Collaborative and Sound Research Designs that Substantiate Best Practice Recommendations for Transgender Health Care. *AMA J Ethics* 2016 Nov 01;18(11):1098-1106 [FREE Full text] [doi: [10.1001/journalofethics.2016.18.11.stas1-1611](https://doi.org/10.1001/journalofethics.2016.18.11.stas1-1611)] [Medline: [27883301](https://pubmed.ncbi.nlm.nih.gov/27883301/)]

Abbreviations

- e-consult:** e-consultation
- EHR:** electronic health record
- ICD:** International Classification of Diseases
- LGBT:** lesbian, gay, bisexual, and transgender
- VA:** Department of Veterans Affairs

Edited by C Lovis; submitted 25.07.18; peer-reviewed by C Liddy, E Keely; comments to author 23.09.18; revised version received 02.10.18; accepted 09.10.18; published 14.01.19.

Please cite as:

*Blosnich JR, Rodriguez KL, Hruska KL, Kavalieratos D, Gordon AJ, Matza A, Mejia SM, Shipherd JC, Kauth MR
Utilization of the Veterans Affairs' Transgender E-consultation Program by Health Care Providers: Mixed-Methods Study
JMIR Med Inform 2019;7(1):e11695*

URL: <http://medinform.jmir.org/2019/1/e11695/>

doi: [10.2196/11695](https://doi.org/10.2196/11695)

PMID: [31344672](https://pubmed.ncbi.nlm.nih.gov/31344672/)

©John R Blosnich, Keri L Rodriguez, Kristina L Hruska, Dio Kavalieratos, Adam J Gordon, Alexis Matza, Susan M Mejia, Jillian C Shipherd, Michael R Kauth. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 14.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Information Technology–Assisted Treatment Planning and Performance Assessment for Severe Thalassemia Care in Low- and Middle-Income Countries: Observational Study

Rajat Kumar Agarwal^{1,2}, BE; Amit Sedai^{1,2}, BE; Kumari Ankita^{1,2}, BE; Lalith Parmar^{1,2}, BE; Rakesh Dhanya², BE; Sunil Dhimal¹, BE; Reshma Srinivas³, MBBS; Ashwini Gowda³, BSc; Pooja Gujjal⁴, MD; Pushpa H^{2,4}, BA; Suman Jain⁵, DCH; J Dasaratha Ramaiah⁶, MD; Sujata Jali⁷, MD; Neelavva Rayappa Tallur⁷, BA; Stalin Ramprakash⁸, MD; Lawrence Faulkner^{2,9}, MD

¹Jagriti InnoHealth Platforms Private Ltd, Bangalore, India

²Sankalp India Foundation, Bangalore, India

³Project Samraksha, Rashtrorathana Parishat, Bangalore, India

⁴Indira Gandhi Institute of Child Health, Bangalore, India

⁵Thalassemia and Sickle Cell Society, Hyderabad, India

⁶Rural Development Trust Hospital, Ananthpur, India

⁷Jawaharlal Nehru Medical College, Belgaum, India

⁸Sankalp-People Tree Centre for Pediatric Bone Marrow Transplant, People Tree Hospitals, Bangalore, India

⁹Cure2Children Foundation, Florence, Italy

Corresponding Author:

Rajat Kumar Agarwal, BE

Jagriti InnoHealth Platforms Private Ltd

84/1A Bannerghatta Main Road

Bangalore, 560076

India

Phone: 91 9880132850

Email: rajat@jagriti.co.in

Abstract

Background: Successful models of information and communication technology (ICT) applied to cost-effective delivery of quality care in low- and middle-income countries (LMIC) are an increasing necessity. Severe thalassemia is one of the most common life-threatening noncommunicable diseases of children globally.

Objective: The aim was to study the impact of ICT on quality of care for severe thalassemia patients in LMIC.

Methods: A total of 1110 patients with severe thalassemia from five centers in India were followed over a 1-year period. The impact of consistent use of a Web-based platform designed to assist comprehensive management of severe thalassemia (ThalCare) on key indicators of quality of care such as minimum (pretransfusion) hemoglobin, serum ferritin, liver size, and spleen size were assessed.

Results: Overall improvements in initial hemoglobin, ferritin, and liver and spleen size were significant ($P < .001$ for each). For four centers, the improvement in mean pretransfusion hemoglobin level was statistically significant ($P < .001$). Four of five centers achieved reduction in mean ferritin levels, with two displaying a significant drop in ferritin ($P = .004$ and $P < .001$). One of the five centers did not record liver and spleen size on palpation, but of the remaining four centers, two witnessed a large drop in liver and spleen size ($P < .01$), one witnessed moderate drop ($P = .05$ for liver; $P = .03$ for spleen size), while the fourth witnessed a moderate increase in liver size ($P = .08$) and insignificant change in spleen size ($P = .12$).

Conclusions: Implementation of computer-assisted treatment planning and performance assessment consistently and positively impacted indexes reflecting effective delivery of care to patients suffering from severe thalassemia in LMIC.

(JMIR Med Inform 2019;7(1):e9291) doi:[10.2196/medinform.9291](https://doi.org/10.2196/medinform.9291)

KEYWORDS

β -thalassemia major; health information technologies; health planning; patient outcome assessment

Introduction

Internet access has grown dramatically in low- and middle-income countries (LMIC) [1]; however, whether improved access to information and communication technologies (ICTs) can improve outcomes is an ongoing debate [2-6]. In fact, the application of information systems to improve access to health care in developing countries has faced many failures related to factors such as mismatch between ICT designs and local user actual needs, system maintenance, upgrades, and repairs [7-9]. The shifting need in many LMIC from infectious diseases and primary health care to chronic noncommunicable diseases (NCD) requiring tertiary care and specialized medical competencies underscores the urgent need to improve access to qualified consultation, patient monitoring, and outcomes assessment. Context-appropriate, friendly, and user-driven ICT systems may play a key role in achieving those goals.

Health care continues to be a challenge in India with only 7 physicians and 17.1 nurses for every 10,000 people against a global mean of 17.1 and 25.3, respectively [10]. Overburdened and underfunded health care prioritizes situations requiring acute medical care. The emergence of NCD in LMIC is driving many families into poverty (an estimated 40 million people a year only in India [11,12]), and the issue of consistency and value assessment in medical care becomes increasingly critical [13]. Last but not least is the need of data management to strengthen health care systems, support governments and other health-providing bodies, and improve research and higher medical education [14,15].

In the pediatric age group, severe hemoglobinopathies, namely severe thalassemia and sickle cell disease (SCD), represent the most frequent life-threatening NCD globally and are a major burden to affected families and health care systems [16-21]. It has been estimated that if the survival rate of children with SCD in Africa increases to only 50% from the current African norm, more than 6 million Africans will be living with SCD [22]. In India, severe thalassemia—a thalassemia syndrome with an inability to spontaneously maintain hemoglobin levels at or above 7 g/dL—is still the largest challenge in the spectrum of hereditary disorders [23-25].

Significant improvements have been made in the care and management of severe thalassemia in the last few decades [26,27]. However, the World Health Organization estimated that only about 9.6% of transfusion-dependent patients suffering from hemoglobin disorders are actually transfused in Southeast Asia [28] and management decisions are often driven by immediate short-term needs rather than established best practices and guidelines. The cumulative impact of years of improper management results in an avoidable morbidity and mortality burden [24,29,30].

The primary objective of care and management of severe thalassemia is to maintain adequate hemoglobin levels while

controlling iron overload [31]. In the context of the developing world, it also includes the need to reverse the impact of ineffective and insufficient management from the past [32,33]. Pretransfusion hemoglobin level is the marker for adequacy of blood transfusions. Serum ferritin is a widely used and acceptable marker for iron overload [34,31]. Liver and/or spleen enlargement also reflect inadequate management and are quite easy to monitor by clinical examination; in fact, liver size on palpation remains one of the most relevant predictors of bone marrow transplant outcome [35,36].

With over 10,000 new patients with severe thalassemia estimated to be born in India alone every year [37], the challenge is to create smart delivery systems that reduce the burden on health care professionals while improving outcomes. We used information technology as an enabler to assist delivery of care to severe thalassemia patients.

Methods

Study Design

We measured the impact of computer-assisted treatment planning and performance assessment in five centers in India that adopted ThalCare between 2011 and 2017, namely Indira Gandhi Institute of Child Health in Bangalore, Project Samraksha (Rashtrathana Parishat) in Bangalore, Thalassemia and Sickle Cell Society in Hyderabad, Rural Development Trust in Ananthpur, and Jai Shivshakti Center for Thalassemia (Jawaharlal Nehru Medical College) in Belgaum.

We selected four quality-of-care indicators: (1) pretransfusion hemoglobin level, (2) serial serum ferritins, (3) liver size, and (4) spleen size. First and fifth quarterly means of each parameter were compared with the intent of quantifying the impact made after a year of use. A total of 4709 visits, 3782 liver measurements, 3825 spleen measurements, and 957 ferritin tests were included in the study. The setup of the centers, staffing, patient age, and gender are summarized in Table 1.

Software Platform Description

The Web-based health ICT platform ThalCare is designed specifically to cater to data management and analytics needs of centers involved in care of severe thalassemia. It was built using free and open-source tools including LAMP (Linux, Apache, MySQL, and PHP) and software stack using Drupal. This cloud-hosted app is accessible through any internet-enabled device with user-specific password-protected accounts. The system was secured and maintained in line with best practices for Web-based health care software. The app was remotely backed up periodically, the software stack was kept updated and guarded against known vulnerabilities, all user activity was logged, and inactivity-triggered timeouts enforced. The users had role-wise filtered access to data related to their own centers.

Table 1. Details of the centers^a and enrollment information.

Category	IGICH	SAM	TSCS	RDT	JSCT
Host institution	Attached to a major academic children's hospital	Attached to a stand-alone blood bank	Attached to a stand-alone blood bank	Attached to a rural hospital	Attached to a medical college
Setup	Government	Private (nonprofit)	Private (nonprofit)	Private (nonprofit)	Private (nonprofit)
Began using ThalCare	Nov 2011	Aug 2013	Nov 2014	Aug 2015	Mar 2016
Doctors, n	1 part-time	1	2	1 part-time	1 part-time
Nurses, n	1	1	2	1	2
Coordinators, n	1	1	2	0	0
Patients, n ^b	154	144	668	88	185
Patient age (years), median ^b	6.5	9.1	6.6	8.4	8.4
Patient sex, n^b					
Male	96	68	58	399	117
Female	58	76	30	269	68

^aIGICH: Indira Gandhi Institute of Child Health; JSCT: Jai Shivshakti Center for Thalassemia; RDT: Rural Development Trust; SAM: Project Samraksha; TSCS: Thalassemia and Sickle Cell Society.

^bData taken at the end of the fifth quarter of adoption of ThalCare.

The system allows capturing patient's contact details, demographic details, socioeconomic information, past severe thalassemia management history, vaccination status, related medical details of other members of the family, etc, at enrollment. Thereafter every patient intervention, including vital signs recordings, clinical notes, blood transfusions, medical prescriptions, laboratory investigations, complications and their management, and growth-related information, are captured on the ICT platform. With comprehensive use of technology, the units are essentially paperless in their operations.

The majority of nurses and coordinators who entered most of the data had never used computer systems before and started using the system after one day of training. In the initial stages, data were reviewed daily and feedback was provided by phone or online sessions. A technology support team was always available on call and provided timely assistance. Within the first month of using the system, all relevant medical data were available on the system.

Scheduling

Requests are sent to the blood bank specifying when blood will be needed, and visits are scheduled sufficiently in advance to match blood availability thus minimizing idle time for patients. An alert is generated for a missed appointment prompting the staff to contact families and reschedule visits.

Treatment Planning

When a patient comes for review, the caregiver updates the clinical record and reviews the system records for alerts on required interventions. The staff plan the day based on the inputs from patient indicators (Textbox 1). The system calculates the volume of blood to be transfused for each child based on preprogrammed configurable formulas; it enables investigations to be done at specified intervals thus facilitating early detection of complications, generates laboratory investigation forms and labels, and interfaces with multiple external laboratory information systems to place requests and retrieve reports.

Textbox 1. Indicators available on the information and communication technology platform.

Biochemistry: displays parameters outside reference range and reminds of the need to repeat the test
Bone marrow transplant eligibility: checks if one or more healthy sibling is available and if age is less than prescribed cut-off age
Complete blood count: warns of cytopenia
Chelation: notifies if chelation information is not available and indicates if there is a need to start chelation; also suggests if dose of the chelator is outside the prescribed ranges
Ferritin: alerts on the grading of the ferritin levels and give reminders for tests
Hemoglobin: alerts if the patient has been discharged with the inappropriate hemoglobin levels in the past visits
Serology: alerts if the patient has one or more transfusion-transmitted infections and reminds for repeating the tests
Transfusion: alerts on the rate of fall of hemoglobin per week and warns of ineffective blood transfusion therapy
Ultrasound: alerts on hepatosplenomegaly as per grading and reminds to repeat the tests
Vaccination: alerts on vaccination status

Care Management

When meeting a patient, the caregiver has a summary of alerts and the system suggests possible interventions. Centers can deactivate certain alerts and only those which require action are visible to avoid alert fatigue [2,38]. Categorized clinical notes allows doctors to find context-appropriate details. Continuous dose adjustment is critical to the success of iron chelation therapy [39,40]: the system prompts if the medication doses do not agree with treatment protocols based on laboratory results and patient's weight. The system also tracks compliance to treatment by tracking the purchase/issue of drugs [41] dispensed directly by centers.

Change in patient's growth, blood counts, and other laboratory values are tracked over time. The system automatically converts height, weight [42,43], and ultrasound measures [44] into age-adjusted z scores and presents them as charts for better visualization. Laboratory values are color-coded using the red-amber-green methodology, and tooltips provide reference ranges. Subsequent patient visits are also scheduled based on the volume of blood transfused and the estimated postransfusion hemoglobin levels.

All complications are recorded and tracked on the system enabling a comprehensive understanding of patient-specific medical needs.

Networking and Collaboration

The ICT platform enables point-of-care professionals to seek advice from more experienced centers. Queries posted on the system are answered asynchronously by experts who are part of the network and have access to the entire clinical history. This enables relatively junior local doctors to team up with specialists and enhance clinical management. Periodic review of the center's progress and challenges is also done through online meetings.

Blood Bank Coordination

The ICT platform allows the clinical team to monitor qualitative and quantitative aspects of associated blood bank support including time elapsed between collection and issue of blood, leukoreduction/depletion, blood request processing time, unavailability of blood locally, demand for replacement blood, and transfusion reaction.

Outcomes and Performance Monitoring

In addition to severe thalassemia treatment planning and daily workflow organization, the ICT platform can generate outcome

analysis and drive quality control. The Thal Report Card is a periodic autogenerated ICT report which summarizes the status of the patients at the center and allows for a quick review of the center's performance. The report is automatically emailed to the health care providers and administrators every week and at the beginning of the month so that everyone involved has a clear view of the status and challenges.

Adverse changes in metrics in the report card act as triggers for quality improvement at the center. Objective evaluation allows discussions to be based on quantified hard data. Most centers have weekly meetings to discuss unit-specific issues and monthly meetings to discuss overall direction and progress. Transparent reporting is also a motivator for involved personnel, who get immediate visibility and recognition, and has allowed nursing staff at these institutions to take on a more active role in patient care and monitoring. The centers also participate in annual preview with peer benchmarking and sharing of best practices.

Transparency, Accountability, and Community Engagement

Another key aspect of the Thal Report Card is that it enables data-driven communication with funding agencies, regulators, administrators, and patient groups to monitor progress. Positive changes act as a catalyst to further aid resource mobilization, while transparent reporting builds accountability leading to continuous quality improvement.

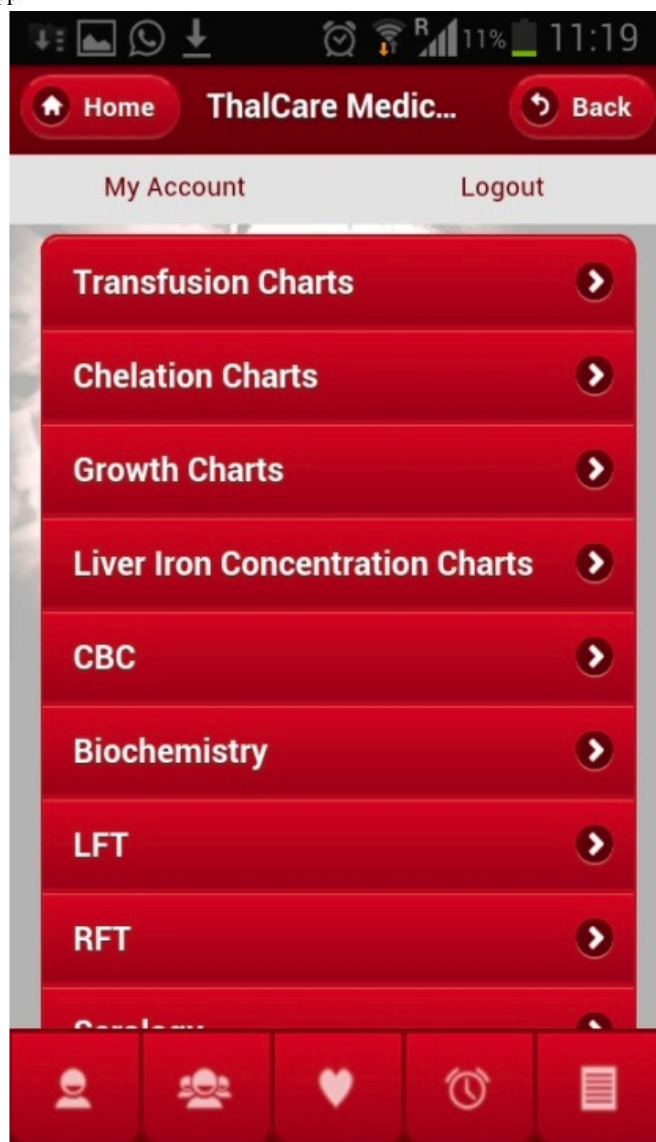
Patient App

The ICT platform allows patients to log in and track their medical progress, see laboratory reports and appointments, and interact with the treating team leading to enhanced awareness and involvement (Figure 1). This helps maintain accurate and updated personal medical information while eliminating the need for paper documents and vastly simplifying record keeping. However, this was useful only for the more educated patients/families.

Statistical Analysis

Data collected from the ICT platform was summarized using MS Excel. The mean values of the pretransfusion hemoglobin levels and liver and spleen sizes were compared using Welch t test (independent two-sample, two-tailed assuming unequal size and unequal variance) and ferritin was compared using matched-pair t tests using R version 3.3.2.

Figure 1. Screenshot of the patient app.



Results

Adequacy of Blood Transfusions

Overall improvement in pretransfusion hemoglobin was highly significant ($P < .001$) and the same was seen in four centers (Table 2). The improvement was statistically significant in four ($P < .001$), whereas the fifth center saw a marginal drop of 0.1 gm/dL ($P = .09$).

The number of visits increased from a mean of 0.7 (range 0.5-1.5) visits per month to 1.1 (range 0.8-1.4) visits per month. Four centers sourced all blood from the attached blood banks. The need for patients to get blood from outside fell from 17% to 11% in the fifth center. The mean duration between collection and issue of blood was 4.3 (SD 2.9) days. In three centers, blood was processed within 3 hours; one center did not measure this time and the remaining center took 4.7 hours. Transfusion-transmitted infection prevalence was 1.8% (range 0%-2.9%). Of this, 1.7% was detected at registration with ThalCare and 0.12% was detected subsequently.

Iron Overload Management

The overall reduction in mean ferritin levels was significant ($P < .001$). Two centers achieved significant reduction in mean ferritin levels ($P < .001$ and $P = .004$), two had an insignificant drop ($P = .80$ and $P = .50$), and the remaining one had an insignificant increase ($P = .80$) (see Table 2).

Of the 983 changes made to chelation drug doses, 521 (53.0%) were made as a response to system-generated alerts for dose change.

Response to Liver and Spleen Size on Palpation

Overall, there was highly significant improvement in liver ($P < .001$) and spleen ($P < .001$) sizes. One center did not record liver and spleen size. Two witnessed highly significant drops in liver and spleen size ($P < .01$), one witnessed moderate drop ($P = .05$ for liver; $P = .03$ for spleen), while the fourth witnessed increase in liver size ($P = .08$) and insignificant change in spleen size ($P = .12$) (see Table 2).

Table 2. Change in the key parameters in the first year of adoption of ThalCare.

Group	Pretransfusion Hb (g/dL)			Ferritin (ng/mL)			Liver size (cm)			Spleen size (cm)		
	Mean (SD)	n	P value	Mean (SD)	n	P value	Mean (SD)	n	P value	Mean (SD)	n	P value
Quarter 1												
Overall	8.0 (1.6)	1009	— ^a	4940 (4628)	427	—	2.4 (2.4)	869	—	3.5 (3.8)	863	—
IGICH ^b	7.1 (1.4)	151	—	3856 (3050)	95	—	3.1 (2.6)	154	—	3.1 (3.3)	154	—
SAM ^c	8.2 (2.0)	232	—	5376 (2843)	99	—	3.8 (2.4)	161	—	5.6 (5.4)	161	—
TSCS ^d	8.2 (1.3)	405	—	3281 (1537)	42	—	1.2 (1.3)	364	—	2.4 (2.8)	361	—
RDT ^e	7.9 (1.6)	221	—	3959 (4051)	29	—	2.9 (2.7)	190	—	4.1 (3.7)	187	—
JSCT ^f	7.5 (1.7)	369	—	5916 (6380)	162	—	—	—	—	—	—	—
Quarter 5												
Overall	8.3 (1.3)	2891	<.001	3517 (2262)	530	<.001	1.6 (2.0)	2802	<.001	2.2 (3.0)	2798	<.001
IGICH	7.7 (1.4)	214	<.001	2849 (1647)	151	.004	0.9 (1.9)	245	<.001	0.8 (1.9)	245	<.001
SAM	8.7 (1.3)	374	<.001	4569 (2916)	130	.80	3.2 (1.9)	345	.01	2.3 (3.5)	339	<.001
TSCS	8.1 (1.3)	1985	.09	3363 (2039)	88	.80	1.3 (1.7)	1886	.08	2.2 (2.9)	1891	.12
RDT	9.1 (1.3)	318	<.001	3404 (2328)	54	.50	2.4 (2.6)	326	.05	3.4 (3.2)	326	.03
JSCT	9.5 (1.4)	440	<.001	3363 (1785)	107	<.001	—	—	—	—	—	—

^aNot applicable.

^bIGICH: Indira Gandhi Institute of Child Health.

^cSAM: Project Samraksha.

^dTSCS: Thalassemia and Sickle Cell Society.

^eRDT: Rural Development Trust.

^fJSCT: Jai Shivshakti Center for Thalassemia.

Other Changes

For each day of work at the thalassemia clinic, a mean 1.3 (range 1.0-1.7) users logged in with at least one record posted from within the center. Multiple log-ins/posts by the same user were counted as 1. The mean records posted per day was 30.6 (range 17.0-53.8).

The total percentage of external consultations ranged from 0% to 92% with a mean of 19.24% (786/4086). Nurses and coordinators from three centers heavily relied on remote advice. Involvement of remote hematologists helped identify 23 patients with SCD and 7 patients with hereditary spherocytosis, all of whom were being inappropriately treated with blood transfusions. This was subsequently corrected.

No nurse initially delegated to work in the thalassemia day care center changed job location during the study period.

Identification of candidates suitable for transplantation led to 391 patients getting human leukocyte antigen-typed and prepared for bone marrow transplantation from these centers. In all, 43 were transplanted with overall survival of 93% and disease-free survival of 77% at a mean cost of US \$13,000 per transplant, including pretransplant preparation and late complications [45,46]. For these patients, online software specific for bone marrow transplantation was used [47].

Discussion

The adoption of an electronic system for patient care was smooth in all centers. The high degree of daily usage (1.3 user log-ins per center and 30.6 posts per day) demonstrates that the ICT platform was well received and operated by the users. We believe that the use of information technology and outcome measurement led to much-needed recognition of the impact created by individuals in the team—especially at a junior level—and may have contributed to the low professional turnover observed in our centers [48].

The ICT platform integrates with the laboratory information systems relieving both health care professionals and parents from the effort of tracking and securing reports. Autogenerated labels for samples, computer-generated blood request forms, and electronic laboratory integration decreased work burden. No additional manpower was recruited just for data management and no additional compensation paid for ICT platform usage. More patients were managed (increase of 0% to 83% with a mean of 18% between centers) with the same staff after the adoption of ICT platform. The cost of using ThalCare was less than US \$200 per month per center.

Of all amendments made to chelation doses, 53% were triggered by the ICT platform. Unintended omissions of clinical interventions such as timely dose modification were significantly reduced, even as the clinicians devoted a similar amount of

time. We observed better adherence to prescribed regular investigations and better estimation of the blood volume to be transfused.

Ensuring regular supply of blood is a considerable challenge in LMIC and often the family is made responsible to organize blood [49]. The ICT platform helped improve the relationship between treatment center and blood banks as pointed out by the reduced need for third-party blood banks. Regular sharing of data on the turnaround time for blood processing, time between collection and issue of blood products, reporting of transfusion reactions, and careful monitoring of transfusion-transmitted infections allowed the blood banks to implement internal changes meeting the specific requirements of the severe thalassemia clinic.

Often the only medical record available for a child diagnosed with severe thalassemia continues to be notebooks, files, or at best a preprinted diary, which are often soiled, worn out, and prone to being lost. The ICT platform relieves the family from the burden of having to maintain the essential medical history of the child. Unfortunately, so far the uptake of the patient app has been very limited primarily for socioeconomic reasons. This is an area of concern which we continue to work on.

Poverty, ignorance, and illiteracy continue to be significant challenges in LMIC [12,50]. Often the responsibility of adhering to treatment protocol, including getting regular tests done and seeking appointments, is passed on to the family. The socially better-placed families end up getting better quality of care—even as the neglect of the most vulnerable ones continues [51]. Systematic monitoring of each patient's clinical status allows early identification of those patients who need more attention—decoupling the family's socioeconomic status from the actual awareness and recognition of the patient's needs. By achieving this, the ICT platform has the potential to improve equity in access to care.

All blood components were provided free of charge [52] and centers were largely supported by nonprofit organizations mentioned in the Acknowledgment section. The mean overall cost per child was about US \$1000 per year. This is a third of the cost of management estimated in India in 2008 (which is likely to have increased with inflation) [53].

With NCD-related health care costs rising and becoming an increasing concern globally, the need for judicious, accountable, and traceable use of resources seems increasingly relevant. The ICT platform enables measurability of the delivery of care, provides a way to maintain an internal check, and to generate outcome benchmarking.

An example of the impact on strategic planning aided by the ICT platform is the decision on purchasing deferoxamine infusion pumps and loaning them to the patients. Similarly, data continues to support the centers in their decision to select the method for leukoreduction of blood products.

The role of ICT in enabling tertiary care in developing countries has been described earlier [47] and challenges on the road to sustainable implementation have been noted [54-56]. Our experience shows successful adoption and sustained use of the ICT platform across five centers in different settings with enhanced outcomes.

Given the general shortage of qualified health care professionals, the ICT platform may have a critical role enabling networking and collaboration. Electronically maintained, well-structured records help overcome wide gaps in the availability of specialists. It is notable that 19% of all consultations were made by health service providers outside the center; the number was as high as 92% for a less experienced center. It is important to highlight that when aided online, the less experienced centers performed as well as their more experienced counterparts. The fact that mismanaged patients were identified and their treatment course corrected by peers from more experienced centers is an example of the possibilities such a model of collaboration brings in. The system can enable that knowledge and best practices are propagated seamlessly allowing centers to be involved in more frequent and well-directed knowledge exchange with peers.

A unique differentiator of this experience has been the fact that unlike the usual practice of limiting the use of ICT to enable collaboration between centers for seeking advice on telemedicine basis or specific periodic interactions, ICT was used for asynchronous teamwork. Apart from being more traceable, this makes it convenient across individual schedules and time zones, thereby creating a sort of patient-specific forum.

In conclusion, our model suggests that computer-assisted treatment planning and performance assessment can significantly improve indexes associated with the effective delivery of care to patients suffering from severe thalassemia. In our experience, a focused context-appropriate, user-driven online IT (Information Technology) tool can have major impact on health care delivery and, importantly, this can occur independently of additional financial and professional resources. Finally, our observation underscores how ICT assists objective outcome reporting, which is the ultimate indicator of quality of care.

Acknowledgments

The authors thank the patients, their parents, and all the nurses and physicians at collaborating institutions. This work was supported by Sankalp India Foundation, Bangalore-India; Jai Shivshakti Health and Education Foundation, Bangalore, India; Cure2Children Foundation, Italy; and Jagriti Innovations, Bangalore, India. We are also grateful to Dr Vikramjit Kanwar for reviewing the manuscript.

Authors' Contributions

RKA designed and performed the research, collected the data, performed the statistical analyses, interpreted the data, and wrote the manuscript; AS, KA LP, RD, SD, and LF designed the research, coordinated patient registration, data management, and data collection; LF, RS, PG, PH, AG, SJ, NRT, JDR, SJ, and SR participated in patient management; all authors critically reviewed and approved the manuscript.

Conflicts of Interest

None declared.

References

1. Poushter J. Pew Research Center. 2016 Feb 22. Smartphone ownership and internet usage continues to climb in emerging economies URL: <http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed 2018-11-23] [WebCite Cache ID 749qqAIbZ]
2. Park H. Health informatics in developing countries: a review of unintended consequences of IT implementations, as they affect patient safety and recommendations on how to address them. *Yearb Med Inform* 2016 Nov 10(1):1-2 [FREE Full text] [doi: [10.15265/IY-2016-028](https://doi.org/10.15265/IY-2016-028)] [Medline: [27830233](https://pubmed.ncbi.nlm.nih.gov/27830233/)]
3. Chavannes NH, Du Puy RS, Bai C. Suggestions for health information technology trials for respiratory disorders in low- and middle-income country settings: what can we learn from trials in high-income country settings? *NPJ Prim Care Respir Med* 2015 Jul 09;25:15045 [FREE Full text] [doi: [10.1038/npjpcrm.2015.45](https://doi.org/10.1038/npjpcrm.2015.45)] [Medline: [26158263](https://pubmed.ncbi.nlm.nih.gov/26158263/)]
4. Kahouei M, Zadeh JM, Roghani PS. The evaluation of the compatibility of electronic patient record (EPR) system with nurses' management needs in a developing country. *Int J Med Inform* 2015 Apr;84(4):263-270. [doi: [10.1016/j.ijmedinf.2014.12.006](https://doi.org/10.1016/j.ijmedinf.2014.12.006)] [Medline: [25618609](https://pubmed.ncbi.nlm.nih.gov/25618609/)]
5. Agarwal S, Perry HB, Long L, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. *Trop Med Int Health* 2015 Aug;20(8):1003-1014 [FREE Full text] [doi: [10.1111/tmi.12525](https://doi.org/10.1111/tmi.12525)] [Medline: [25881735](https://pubmed.ncbi.nlm.nih.gov/25881735/)]
6. Huckvale C, Car J, Akiyama M, Jaafar S, Khoja T, Bin KA, et al. Information technology for patient safety. *Qual Saf Health Care* 2010 Aug;19 Suppl 2:i25-i33 [FREE Full text] [doi: [10.1136/qshc.2009.038497](https://doi.org/10.1136/qshc.2009.038497)] [Medline: [20693213](https://pubmed.ncbi.nlm.nih.gov/20693213/)]
7. Heeks R. Information systems and developing countries: failure, success, and local improvisations. *Inf Soc* 2002 Mar;18(2):101-112. [doi: [10.1080/01972240290075039](https://doi.org/10.1080/01972240290075039)]
8. Moucheraud C, Schwitters A, Boudreaux C, Giles D, Kilmarx PH, Ntolo N, et al. Sustainability of health information systems: a three-country qualitative study in southern Africa. *BMC Health Serv Res* 2017 Jan 10;17(1):23 [FREE Full text] [doi: [10.1186/s12913-016-1971-8](https://doi.org/10.1186/s12913-016-1971-8)] [Medline: [28073361](https://pubmed.ncbi.nlm.nih.gov/28073361/)]
9. Irinoye OO, Ayandiran EO, Fakunle I, Mtshali N. Nurses' perception and barriers to use of information communication technology in a teaching hospital in Nigeria. *Comput Inform Nurs* 2013 Aug;31(8):394-400. [doi: [10.1097/NXN.0b013e3182997a6e](https://doi.org/10.1097/NXN.0b013e3182997a6e)] [Medline: [23851710](https://pubmed.ncbi.nlm.nih.gov/23851710/)]
10. World Health Statistics 2015. Geneva: World Health Organization; 2015. URL: <http://public.eblib.com/choice/publicfullrecord.aspx?p=2070501ISBN:978-92-4-069443-9> [accessed 2018-11-23] [WebCite Cache ID 749rDP59y]
11. Overdorf J. Public Radio International. 2012 Oct 17. India: medical expenses plunge 40 million people into poverty each year URL: <https://www.pri.org/stories/2012-10-17/india-medical-expenses-plunge-40-million-people-poverty-each-year> [accessed 2018-11-23] [WebCite Cache ID 749rGDVMr]
12. Balarajan Y, Selvaraj S, Subramanian S. Health care and equity in India. *Lancet* 2011 Feb;377(9764):505-515. [doi: [10.1016/S0140-6736\(10\)61894-6](https://doi.org/10.1016/S0140-6736(10)61894-6)]
13. Iskrov G, Stefanov R. Ensuring transparency and consistency in the value assessment of cancer therapies. *J Clin Oncol* 2016 Aug 20;34(24):2938-2939. [doi: [10.1200/JCO.2015.63.6076](https://doi.org/10.1200/JCO.2015.63.6076)] [Medline: [27298423](https://pubmed.ncbi.nlm.nih.gov/27298423/)]
14. Elzawayy AM. Could African and low- and middle-income countries contribute scientifically to global cancer care? *J Glob Oncol* 2015 Dec;1(2):49-53 [FREE Full text] [doi: [10.1200/JGO.2015.001032](https://doi.org/10.1200/JGO.2015.001032)] [Medline: [28804772](https://pubmed.ncbi.nlm.nih.gov/28804772/)]
15. The Lancet. Health in India, 2017. *Lancet* 2017 Jan;389(10065):127. [doi: [10.1016/S0140-6736\(17\)30075-2](https://doi.org/10.1016/S0140-6736(17)30075-2)]
16. de Silva S, Fisher CA, Premawardhena A, Lamabadusuriya SP, Peto TE, Perera G, et al. Thalassaemia in Sri Lanka: implications for the future health burden of Asian populations. Sri Lanka Thalassaemia Study Group. *Lancet* 2000 Mar 04;355(9206):786-791. [Medline: [10711926](https://pubmed.ncbi.nlm.nih.gov/10711926/)]
17. Riewpaiboon A, Nuchprayoon I, Torcharus K, Indaratna K, Thavorncharoensap M, Ubol B. Economic burden of beta-thalassemia/Hb E and beta-thalassemia major in Thai children. *BMC Res Notes* 2010 Jan 30;3:29 [FREE Full text] [doi: [10.1186/1756-0500-3-29](https://doi.org/10.1186/1756-0500-3-29)] [Medline: [20181056](https://pubmed.ncbi.nlm.nih.gov/20181056/)]
18. Scalone L, Mantovani LG, Krol M, Rofail D, Ravera S, Bisconte MG, et al. Costs, quality of life, treatment satisfaction and compliance in patients with beta-thalassemia major undergoing iron chelation therapy: the ITHACA study. *Curr Med Res Opin* 2008 Jul;24(7):1905-1917. [doi: [10.1185/03007990802160834](https://doi.org/10.1185/03007990802160834)] [Medline: [18507891](https://pubmed.ncbi.nlm.nih.gov/18507891/)]

19. Payne KA, Rofail D, Baladi J, Viala M, Abetz L, Desrosiers M, et al. Iron chelation therapy: clinical effectiveness, economic burden and quality of life in patients with iron overload. *Adv Ther* 2008 Aug;25(8):725-742. [doi: [10.1007/s12325-008-0085-z](https://doi.org/10.1007/s12325-008-0085-z)] [Medline: [18704280](https://pubmed.ncbi.nlm.nih.gov/18704280/)]
20. Weidlich D, Kefalas P, Guest JF. Healthcare costs and outcomes of managing β -thalassaemia major over 50 years in the United Kingdom. *Transfusion* 2016 May;56(5):1038-1045. [doi: [10.1111/trf.13513](https://doi.org/10.1111/trf.13513)] [Medline: [27041389](https://pubmed.ncbi.nlm.nih.gov/27041389/)]
21. Ho W, Lin K, Wang J, Hwang J, Chung C, Lin D, et al. Financial burden of national health insurance for treating patients with transfusion-dependent thalassaemia in Taiwan. *Bone Marrow Transplant* 2006 Mar;37(6):569-574. [doi: [10.1038/sj.bmt.1705287](https://doi.org/10.1038/sj.bmt.1705287)] [Medline: [16462756](https://pubmed.ncbi.nlm.nih.gov/16462756/)]
22. Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. *Blood* 2010 Jun 03;115(22):4331-4336 [FREE Full text] [doi: [10.1182/blood-2010-01-251348](https://doi.org/10.1182/blood-2010-01-251348)] [Medline: [20233970](https://pubmed.ncbi.nlm.nih.gov/20233970/)]
23. Mohanty D, Colah RB, Gorakshakar AC, Patel RZ, Master DC, Mahanta J, et al. Prevalence of β -thalassaemia and other haemoglobinopathies in six cities in India: a multicentre study. *J Community Genet* 2013 Jan;4(1):33-42 [FREE Full text] [doi: [10.1007/s12687-012-0114-0](https://doi.org/10.1007/s12687-012-0114-0)] [Medline: [23086467](https://pubmed.ncbi.nlm.nih.gov/23086467/)]
24. Agarwal MB. Advances in management of thalassaemia. *Indian J Pediatr* 2009 Feb;76(2):177-184. [doi: [10.1007/s12098-009-0048-7](https://doi.org/10.1007/s12098-009-0048-7)] [Medline: [19330306](https://pubmed.ncbi.nlm.nih.gov/19330306/)]
25. Weatherall DJ. The challenge of haemoglobinopathies in resource-poor countries. *Br J Haematol* 2011 Sep;154(6):736-744. [doi: [10.1111/j.1365-2141.2011.08742.x](https://doi.org/10.1111/j.1365-2141.2011.08742.x)] [Medline: [21726207](https://pubmed.ncbi.nlm.nih.gov/21726207/)]
26. Galanello R, Origa R. Beta-thalassaemia. *Orphanet J Rare Dis* 2010 May 21;5:11 [FREE Full text] [doi: [10.1186/1750-1172-5-11](https://doi.org/10.1186/1750-1172-5-11)] [Medline: [20492708](https://pubmed.ncbi.nlm.nih.gov/20492708/)]
27. Rund D. Thalassaemia 2016: modern medicine battles an ancient disease. *Am J Hematol* 2016 Jan;91(1):15-21 [FREE Full text] [doi: [10.1002/ajh.24231](https://doi.org/10.1002/ajh.24231)] [Medline: [26537527](https://pubmed.ncbi.nlm.nih.gov/26537527/)]
28. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 2008 Jun;86(6):480-487 [FREE Full text] [Medline: [18568278](https://pubmed.ncbi.nlm.nih.gov/18568278/)]
29. Prakash A, Aggarwal R. Thalassaemia major in adults: short stature, hyperpigmentation, inadequate chelation, and transfusion-transmitted infections are key features. *N Am J Med Sci* 2012 Mar;4(3):141-144 [FREE Full text] [doi: [10.4103/1947-2714.93886](https://doi.org/10.4103/1947-2714.93886)] [Medline: [22454828](https://pubmed.ncbi.nlm.nih.gov/22454828/)]
30. Verma IC, Saxena R, Kohli S. Past, present & future scenario of thalassaemic care & control in India. *Indian J Med Res* 2011 Oct;134:507-521 [FREE Full text] [Medline: [22089615](https://pubmed.ncbi.nlm.nih.gov/22089615/)]
31. Cappellini M, Cohen A, Porter J, Taher A, Viprakasit V. Guidelines for the management of transfusion dependent thalassaemia (TDT). Nicosia, Cyprus: Thalassaemia International Federation; 2014. URL: <https://www.ncbi.nlm.nih.gov/books/NBK269382/> [accessed 2018-11-28] [WebCite Cache ID 74HA1Adwb]
32. Yin X, Wu Z, He Y, Zhou T, Zhou Y, Zhang XH. Treatment and complications of thalassaemia major in Guangxi, Southern China. *Pediatr Blood Cancer* 2011 Dec 15;57(7):1174-1178. [doi: [10.1002/pbc.23101](https://doi.org/10.1002/pbc.23101)] [Medline: [21394896](https://pubmed.ncbi.nlm.nih.gov/21394896/)]
33. Hashemizadeh H, Noori R, Kolagari S. Assessment hepatomegaly and liver enzymes in 100 Patients with beta thalassaemia major in Mashhad, Iran. *Iran J Ped Hematol Oncol* 2012;2(4):171-177 [FREE Full text] [Medline: [24575259](https://pubmed.ncbi.nlm.nih.gov/24575259/)]
34. Olivieri NF, Brittenham GM. Management of the thalassaemias. *Cold Spring Harb Perspect Med* 2013 Jun 01;3(6) [FREE Full text] [doi: [10.1101/cshperspect.a011767](https://doi.org/10.1101/cshperspect.a011767)] [Medline: [23732853](https://pubmed.ncbi.nlm.nih.gov/23732853/)]
35. Lucarelli G, Galimberti M, Polchi P, Angelucci E, Baronciani D, Giardini C, et al. Bone marrow transplantation in patients with thalassaemia. *N Engl J Med* 1990 Feb 15;322(7):417-421. [doi: [10.1056/NEJM199002153220701](https://doi.org/10.1056/NEJM199002153220701)] [Medline: [2300104](https://pubmed.ncbi.nlm.nih.gov/2300104/)]
36. Mathews V, George B, Deotare U, Lakshmi K, Viswabandya A, Daniel D, et al. A new stratification strategy that identifies a subset of class III patients with an adverse prognosis among children with beta thalassaemia major undergoing a matched related allogeneic stem cell transplantation. *Biol Blood Marrow Transplant* 2007 Aug;13(8):889-894 [FREE Full text] [doi: [10.1016/j.bbmt.2007.05.004](https://doi.org/10.1016/j.bbmt.2007.05.004)] [Medline: [17640592](https://pubmed.ncbi.nlm.nih.gov/17640592/)]
37. Thacker N. Prevention of thalassaemia in India. *Indian Pediatr* 2007 Sep;44(9):647-648 [FREE Full text] [Medline: [17921552](https://pubmed.ncbi.nlm.nih.gov/17921552/)]
38. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017 Apr 10;17(1):36 [FREE Full text] [doi: [10.1186/s12911-017-0430-8](https://doi.org/10.1186/s12911-017-0430-8)] [Medline: [28395667](https://pubmed.ncbi.nlm.nih.gov/28395667/)]
39. Aydinok Y, Kattamis A, Viprakasit V. Current approach to iron chelation in children. *Br J Haematol* 2014 Jun;165(6):745-755. [doi: [10.1111/bjh.12825](https://doi.org/10.1111/bjh.12825)] [Medline: [24646011](https://pubmed.ncbi.nlm.nih.gov/24646011/)]
40. Taher A, El-Beshlawy A, Elalfy MS, Al Zir K, Daar S, Habr D, et al. Efficacy and safety of deferasirox, an oral iron chelator, in heavily iron-overloaded patients with beta-thalassaemia: the ESCALATOR study. *Eur J Haematol* 2009 Jun;82(6):458-465 [FREE Full text] [doi: [10.1111/j.1600-0609.2009.01228.x](https://doi.org/10.1111/j.1600-0609.2009.01228.x)] [Medline: [19187278](https://pubmed.ncbi.nlm.nih.gov/19187278/)]
41. Mishra AK, Tiwari A. Iron overload in beta thalassaemia major and intermedia patients. *Maedica (Buchar)* 2013 Sep;8(4):328-332 [FREE Full text] [Medline: [24790662](https://pubmed.ncbi.nlm.nih.gov/24790662/)]
42. World Health Organization. The WHO child growth standards (0-5 years) URL: <http://www.who.int/childgrowth/standards/en/> [accessed 2018-11-23] [WebCite Cache ID 749sAMHY4]
43. World Health Organization. Geneva: WHO Growth reference data for 5-19 years URL: <http://www.who.int/growthref/en/> [accessed 2018-11-23] [WebCite Cache ID 749sIkTiQ]

44. Dhingra B, Sharma S, Mishra D, Kumari R, Pandey RM, Aggarwal S. Normal values of liver and spleen size by ultrasonography in Indian children. *Indian Pediatr* 2010 Jun;47(6):487-492 [FREE Full text] [Medline: [19736366](#)]
45. Faulkner L, Uderzo C, Khalid S, Marwah P, Soni R, Yaqub N, et al. ATG vs thiotepea with busulfan and cyclophosphamide in matched-related bone marrow transplantation for thalassemia. *Blood Adv* 2017 May 23;1(13):792-801 [FREE Full text] [doi: [10.1182/bloodadvances.2016004119](#)] [Medline: [29296723](#)]
46. Agarwal R, Kumari A, Sedai A, Parmar L, Dhanya R, Faulkner L. The case for high resolution extended 6-loci HLA typing for identifying related donors in the Indian subcontinent. *Biol Blood Marrow Transplant* 2017 Sep;23(9):1592-1596 [FREE Full text] [doi: [10.1016/j.bbmt.2017.05.030](#)] [Medline: [28603069](#)]
47. Agarwal RK, Sedai A, Dhimal S, Ankita K, Clemente L, Siddique S, et al. A prospective international cooperative information technology platform built using open-source tools for improving the access to and safety of bone marrow transplantation in low- and middle-income countries. *J Am Med Inform Assoc* 2014;21(6):1125-1128 [FREE Full text] [doi: [10.1136/amiainl-2013-002594](#)] [Medline: [24714444](#)]
48. Bhattacharya I, Ramachandran A. A path analysis study of retention of healthcare professionals in urban India using health information technology. *Hum Resour Health* 2015 Jul 31;13:65 [FREE Full text] [doi: [10.1186/s12960-015-0055-x](#)] [Medline: [26228778](#)]
49. Anand A. BBC News. 2015 Jan 27. Blood for sale: India's illegal "red market" URL: <https://www.bbc.com/news/business-30273994> [accessed 2018-11-23] [WebCite Cache ID 749sa1kxj]
50. Rajan K, Kennedy J, King L. Is wealthier always healthier in poor countries? The health implications of income, inequality, poverty, and literacy in India. *Soc Sci Med* 2013 Jul;88:98-107. [doi: [10.1016/j.socscimed.2013.04.004](#)] [Medline: [23702215](#)]
51. Gwatkin DR, Bhuiya A, Victora CG. Making health systems more equitable. *Lancet* 2004;364(9441):1273-1280. [doi: [10.1016/S0140-6736\(04\)17145-6](#)] [Medline: [15464189](#)]
52. Guidelines for Recovery of Processing Charges for Blood and Blood Components. New Delhi: National Blood Transfusion Council, Government of India; 2014. URL: <http://www.naco.gov.in/sites/default/files/guidelines.pdf> [accessed 2018-11-28] [WebCite Cache ID 74HB4CR1b]
53. Ghosh K, Colah R. Control and Management of Thalassemia and Other Hemoglobinopathies in the Indian Subcontinent : Synoptic Views. Mumbai: National Institute of Immunohaematology; 2008.
54. Luna D, Almerares A, Mayan JC, González BD, Otero C. Health informatics in developing countries: going beyond pilot practices to sustainable implementations: a review of the current challenges. *Healthc Inform Res* 2014 Jan;20(1):3-10 [FREE Full text] [doi: [10.4258/hir.2014.20.1.3](#)] [Medline: [24627813](#)]
55. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. *Bull World Health Organ* 2012 May 1;90(5):332-340 [FREE Full text] [doi: [10.2471/BLT.11.099820](#)] [Medline: [22589566](#)]
56. Kostkova P. Grand challenges in digital health. *Front Public Health* 2015;3:134 [FREE Full text] [doi: [10.3389/fpubh.2015.00134](#)] [Medline: [26000272](#)]

Abbreviations

IGICH: Indira Gandhi Institute of Child Health
JSCT: Jai Shiyshakti Center for Thalassemia
ICT: information and communication technology
LMIC: low- and middle-income countries
NCD: noncommunicable diseases
RDT: Rural Development Trust
SAM: Project Samraksha
SCD: sickle cell disease
TSCS: Thalassemia and Sickle Cell Society

Edited by G Eysenbach; submitted 27.10.17; peer-reviewed by N Chavannes, R Pankomera; comments to author 28.12.17; revised version received 26.01.18; accepted 19.07.18; published 23.01.19.

Please cite as:

Agarwal RK, Sedai A, Ankita K, Parmar L, Dhanya R, Dhimal S, Srinivas R, Gowda A, Gujjal P, H P, Jain S, Ramaiah JD, Jali S, Tallur NR, Ramprakash S, Faulkner L
Information Technology-Assisted Treatment Planning and Performance Assessment for Severe Thalassemia Care in Low- and Middle-Income Countries: Observational Study
JMIR Med Inform 2019;7(1):e9291
URL: <http://medinform.jmir.org/2019/1/e9291/>
doi: [10.2196/medinform.9291](#)
PMID: [30672740](#)

©Rajat Kumar Agarwal, Amit Sedai, Kumari Ankita, Lalith Parmar, Rakesh Dhanya, Sunil Dhimal, Reshma Srinivas, Ashwini Gowda, Pooja Gujjal, Pushpa H, Suman Jain, J Dasaratha Ramaiah, Sujata Jali, Neelavva Rayappa Tallur, Stalin Ramprakash, Lawrence Faulkner. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 23.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A New Insight Into Missing Data in Intensive Care Unit Patient Profiles: Observational Study

Anis Sharafoddini¹, MSc; Joel A Dubin^{1,2}, PhD; David M Maslove³, MD, MS, FRCPC; Joon Lee^{1,4,5}, PhD

¹Health Data Science Lab, School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada

²Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada

³Department of Critical Care Medicine, Queen's University, Kingston, ON, Canada

⁴Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

⁵Department of Cardiac Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:

Joon Lee, PhD

Department of Community Health Sciences

Cumming School of Medicine

University of Calgary

3280 Hospital Drive Northwest

Calgary, AB, T2N 4Z6

Canada

Phone: 1 403 220 2968

Email: joonwu.lee@ucalgary.ca

Abstract

Background: The data missing from patient profiles in intensive care units (ICUs) are substantial and unavoidable. However, this incompleteness is not always random or because of imperfections in the data collection process.

Objective: This study aimed to investigate the potential hidden information in data missing from electronic health records (EHRs) in an ICU and examine whether the presence or missingness of a variable itself can convey information about the patient health status.

Methods: Daily retrieval of laboratory test (LT) measurements from the Medical Information Mart for Intensive Care III database was set as our reference for defining complete patient profiles. *Missingness indicators* were introduced as a way of representing presence or absence of the LTs in a patient profile. Thereafter, various feature selection methods (filter and embedded feature selection methods) were used to examine the predictive power of missingness indicators. Finally, a set of well-known prediction models (logistic regression [LR], decision tree, and random forest) were used to evaluate whether the absence status itself of a variable recording can provide predictive power. We also examined the utility of missingness indicators in improving predictive performance when used with observed laboratory measurements as model input. The outcome of interest was in-hospital mortality and mortality at 30 days after ICU discharge.

Results: Regardless of mortality type or ICU day, more than 40% of the predictors selected by feature selection methods were missingness indicators. Notably, employing missingness indicators as the only predictors achieved reasonable mortality prediction on all days and for all mortality types (for instance, in 30-day mortality prediction with LR, we achieved area under the curve of the receiver operating characteristic [AUROC] of 0.6836 ± 0.012). Including indicators with observed measurements in the prediction models also improved the AUROC; the maximum improvement was 0.0426. Indicators also improved the AUROC for Simplified Acute Physiology Score II model—a well-known ICU severity of illness score—confirming the additive information of the indicators (AUROC of 0.8045 ± 0.0109 for 30-day mortality prediction for LR).

Conclusions: Our study demonstrated that the presence or absence of LT measurements is informative and can be considered a potential predictor of in-hospital and 30-day mortality. The comparative analysis of prediction models also showed statistically significant prediction improvement when indicators were included. Moreover, missing data might reflect the opinions of examining clinicians. Therefore, the absence of measurements can be informative in ICUs and has predictive power beyond the measured data themselves. This initial case study shows promise for more in-depth analysis of missing data and its informativeness in ICUs. Future studies are needed to generalize these results.

(*JMIR Med Inform* 2019;7(1):e11605) doi:[10.2196/11605](https://doi.org/10.2196/11605)

KEYWORDS

electronic health records; clinical laboratory tests; machine learning; hospital mortality

Introduction

Background

The increased adoption of electronic health record (EHR) systems has boosted interest in the secondary use of EHR data [1]. Although the literature has introduced various dimensions for EHR data quality, completeness and correctness have been reported as the fundamental dimensions [1,2]. Although these issues can also be observed in paper-based records, EHR brought us the opportunity to identify them faster and helped us with addressing them. The data missing from clinical contexts are substantial [3,4] and unavoidable [5]; many studies have focused on resolving this issue [6-8]. Although many researchers treat missing data as a challenge [9-18], others continue to debate whether lack of completeness also provides useful information [4,19-21]. Researchers do agree that a part of this incompleteness is not random or because of imperfections in the data collection process [21,22]. Recently, Angiel et al [21] demonstrated that the laboratory ordering time (ie, the interval between 2 orders of a laboratory test; LT) for some LT is more informative than the actual values in predicting 3-year survival. Our study focuses on systematically investigating the implications or possible value of lack of data, particularly in intensive care units (ICUs) and proposes a representation method for missing data to capture hidden information. In general, 2 reasons are given for missing data in EHRs:

- No intention to collect: the clinical variable was never measured because there was no clinical indication to do so—the patient was not suffering from a relevant symptom or comorbidity [4] or it could not be measured [19].
- Intention to collect: records are missing although the variables were measured [4].

Therefore, the health care process (eg, clinicians' decision to order a test and nurse data entry) affects the recorded EHR and can cause incompleteness in data.

Incomplete EHR data can complicate or prohibit the data analysis process, as many machine learning (ML) algorithms assume that there are no missing data in the dataset or require users to clean the data in the preprocessing stage and so provide a complete dataset. Therefore, from a research perspective, the ideal situation is to increase the amount and accuracy of EHR documentation by employing approaches that focus on intention to collect such as reducing the error in data entry or increasing data documentation in terms of resolution. Although the current amount of testing and bloodwork has been reported as actually redundant in ICUs [23-25] and requires extra time and work from clinicians [4], these approaches suffer from their own shortcomings. Besides analytical methods that can handle missing data (that are missing at random) such as decision trees (DTs) or mixed-effects models for longitudinal data, other approaches usually assume missing data are missing completely at random. In general, the literature proposes 3 analytical approaches: complete case analysis (CCA) or deletion, available case analysis (ACA), and imputation.

CCA starts with the list of variables included in the analysis and discards records with missing data on any of the variables. However, this subsample might not be a random sample of the population. Although researchers argue that sample selection based on the predefined eligibility criteria in randomized clinical trials can limit the external generalizability of these studies [26], CCA in studies using EHR data can also potentially threaten the external validity of a study [19] and cause bias as the literature shows a statistically significant relationship between severity of illness and data completeness [20]. A study [19] on 10,000 EHRs from patients receiving anesthetic service showed that patients with an anesthesiologists physical status (ASA) [27] class-4 fitness rating had 5.05 more days with laboratory results and 6.85 more days with medication orders than patients with ASA class 1, suggesting more data are recorded for sicker patients than healthier patients. Thus, imposing complete case requirements when using EHR data for secondary use can cause bias toward selecting patients with more severe conditions (or several comorbidities). Despite this drawback, CCA has been identified as the leading approach in studies on ICU data [28]. That said, CCA provides valid inference only when data are missing completely at random (MCAR), which is unlikely in practice [29].

The ACA (or pairwise deletion) uses all available data for a given analysis. In other words, it maximizes the availability of data by an analysis-by-analysis basis [30]. The advantage of this method is that more data are included in each analysis than with CCA. It also allows for valid inference by likelihood-based models when missing data are ignorable—often the case when the data are missing at random (MAR) [29]. Although ACA is an improvement to CCA [30], it also has limitations. As different samples are being used in each analysis, not only is comparison of various analyses impossible [31] but also using different samples for estimating the parameters of interest has occasionally led to biased or mathematically inconsistent results [32-34].

Imputation methods, which try to draw inferences from incomplete data, rely on knowing the mechanism of missingness, which cannot be validated from the available data. Single imputation methods suffer from 2 problems. First, an inference based on imputed data can be biased if the underlying assumptions are not valid. Second, because imputed data are assumed to be true, the model's statistical precision is overstated. Multiple imputation methods, in spite of their promising performance, rely on parametric assumptions that, if not valid, can lead to incorrect imputation. Due to these limitations, imputation methods should be used with caution and checking underlying assumptions with clinicians is highly recommended [5]. However, Gorelick [35], in a simulation study, demonstrated that either CCA or imputation could cause bias in predictive modeling, and that assuming missing values to be normal when missingness rates are high and substituting them with normal values would also cause substantial bias. In brief, if primary assumptions are not fully satisfied, neither considering complete or available cases nor imputating missing data is likely to yield

reliable results. Furthermore, these statistical methods on their own are not sufficient to capture the hidden information about the patient health status and care process in the complex EHR data. Alternatively, we can try to learn from what is missing rather than only dealing with missingness as a deficiency.

Objectives

This case study provides evidence that missing data in ICU might be missing because of the patient's health status or health care process and introduces a new method for representing patient profiles. In this representation, auxiliary variables, called indicators, are used to represent the presence or absence of a measurement and might convey the possible hidden information in the missing data. Then, by employing various analytical methods, this study attempts to demonstrate the informativeness of missing data. In the rest of the study, the term *missing data* is used to describe not-at-random missing information in patient profiles. In other words, the potential informativeness of data that has not been recorded by choice is of interest.

Methods

Measurement Protocol and Data Collection

As patient monitoring strongly relies on clinical needs, no universal standards for ICU data completeness have been established [36-38]. However, a study by Frassica in 2005 [39] published a list of the top 80% of LTs common to all ICU patients within a university teaching hospital. We revised this list based on the presence of these tests in our database and updated it with input from an ICU clinician to reflect current practices (Textbox 1).

Textbox 1. A total of 36 laboratory tests used in investigating informativeness of missing data.

Variable category and variables

Top 80% laboratory tests and profiles common to all intensive care units [39] reviewed and revised by domain expert

- Alanine aminotransferase (ALT)
- Alkaline phosphatase (ALK)
- Aspartate aminotransferase (AST)
- Arterial blood gases: pH, partial pressure of carbon dioxide (PCO₂), and partial pressure of oxygen (PO₂)
- Base excess (BE)
- Basic metabolic panel: sodium (Na), potassium (K), chloride (Cl), bicarbonate (HCO₃), anion gap (AG), blood glucose (BG), blood urea nitrogen (BUN), and creatinine (Cr)
- Complete blood count: white blood cells (WBCs), red blood cells (RBCs), hemoglobin (HGB), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red cell distribution width (RDW), platelet count (PLT), absolute monocytes (MO), absolute eosinophils (EO), absolute basophils (BA), absolute lymphocytes (LY), and absolute neutrophils (NE)
- Lactate (Lac)
- Calcium (Ca)
- Magnesium (Mg)
- Phosphate (Phos)
- Partial thromboplastin time (PTT)
- Prothrombin time (PT)
- Total bilirubin (TBil)

The data for this study were collected from the Medical Information Mart for Intensive Care III (MIMIC-III) [40] database which contains data from 38,597 distinct adult patients admitted to the Beth Israel Deaconess Medical Center in Boston, Massachusetts, between 2001 and 2012. For patient cohort selection, a tailored version of the generalized cohort selection heuristics for retrospective EHR studies introduced by Harrell et al [41] was used. The data for first admission to 1 of the 5 ICUs—medical ICU, surgical ICU, cardiac care unit, cardiac surgery recovery unit, and trauma surgical ICU—were extracted for adult patients (aged 15 years or older). Included patients must have had at least one data point in any of the variable categories during the first, second, and third days of their ICU stay.

Data Preprocessing and Missing Data Representation

Each day's extracted data were mapped into a matrix with columns for measurements and rows for patients. Therefore, we had a column for each daily measurement of LTs, resulting in 36 columns for LTs. An auxiliary matrix was generated to store binary values reflecting the presence (0) or absence (1) of measurements. As many well-performing ML algorithms are designed to work with a complete data matrix, 2 methods—predictive mean matching (PMM) [42] and hot deck (HD)—were used to impute missing values. PMM is a commonly used and well-accepted imputation method in public health research [43] and is also robust against model misspecification [44]. HD imputation is used commonly in applied data analysis when missing data exist [45].

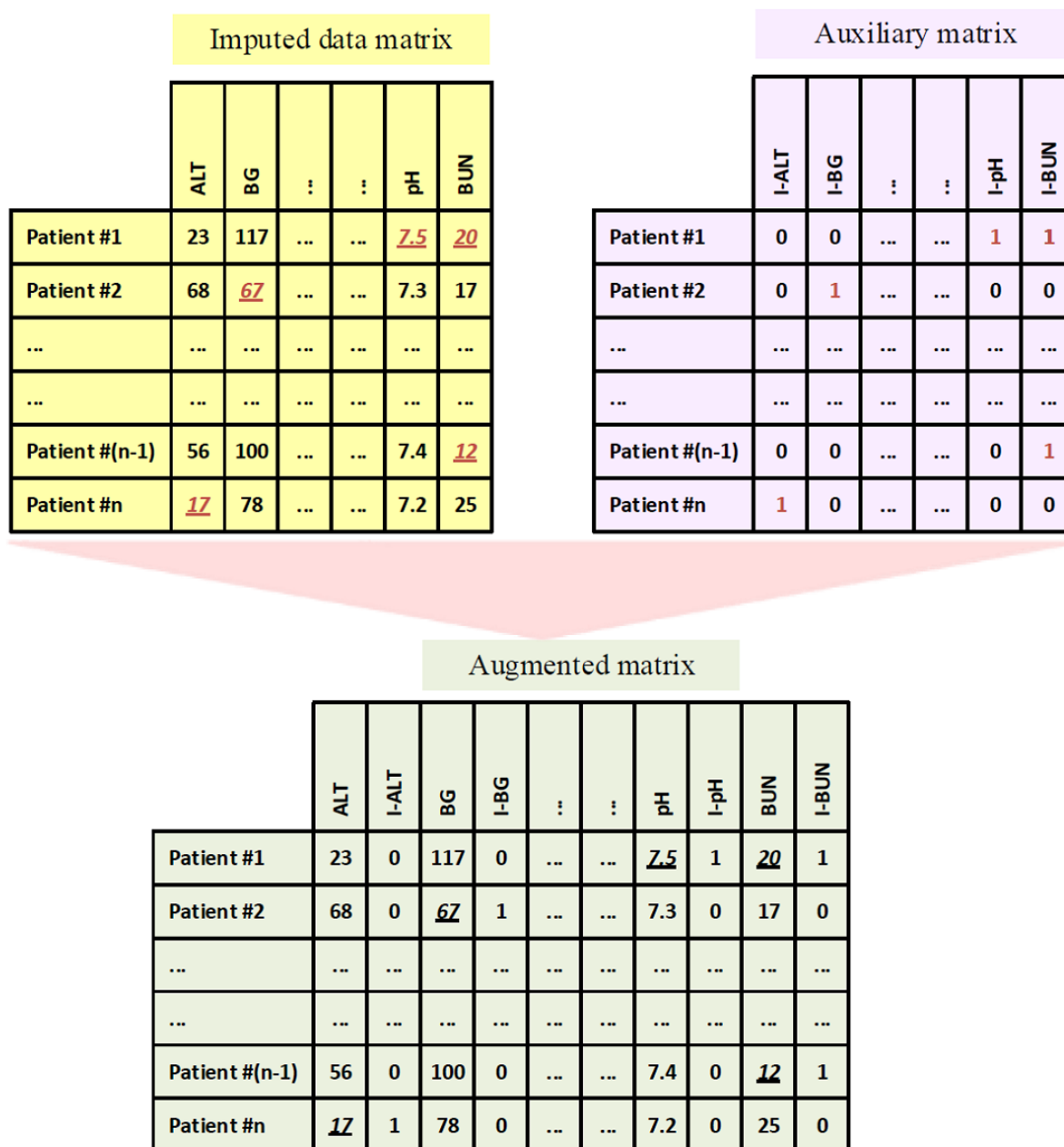
Given that imputed values are indistinguishable to the ML algorithm from true values, we combined the original matrix and auxiliary matrix to form an augmented matrix that directly indicates where values were imputed. This was done to mitigate the risk of treating imputed values the same as actual values, in a setting where the underlying reason for missing data is not fully known (Figure 1). Missing data indicators in this augmented matrix might also provide extra information about the reliability of the values (actual and imputed values) and potentially preserve any meaningful missing data patterns. Missingness indicators have been used as a method of handling missing data in epidemiological and clinical studies. However, in the current use of indicators, missing values are set to a fixed

value (0 or the normal value for the variable) and the indicators are used as dummy variables in analytical models to indicate that a value was missing [46,47]. Studies have shown that this method causes bias as the missing values are imputed with a single value [48]. In our study, we are not using indicators as dummy variables; instead, we are introducing them as a source of information to be used besides imputation methods.

Validation

Several validation techniques are available in medical research. In this study, for all experiments where applicable, we used cross-validation technique (10-fold cross-validation). We also repeated the cross-validation procedure several times (20 times) to acquire more stable results as suggested in the literature [49].

Figure 1. An example of the augmented data matrix, the imputed data matrix (imputed values are underlined and italicized), and the auxiliary matrix (containing the missingness indicators: 0-present, 1-absent).



Assessments

Exploratory Analysis

First, the trends of missingness among LTs were visualized for comparison. Afterward, pairwise correlation among indicators, using Phi coefficient, was done to explore the general behavior of missingness. The Elixhauser [50] and the Charlson [51] comorbidity indices are the most common comorbidity scores in clinical applications. The literature has shown that the Elixhauser Comorbidity Index (ECI) in general has the best performance [52-55]. This better performance can be the result of (1) including new comorbidities in ECI, (2) the differences in the coding of variables common between both indices, or (3) a combination of the first and second factors [53]. The Simplified Acute Physiology Score II (SAPS-II) [56] scoring system that has been widely used by most ICUs for predicting illness severity was also chosen. Therefore, the association of missingness rates with ECI and SAPS-II was investigated using Spearman correlation. Besides the clinical information, SAPS-II also has the information about type of admission (scheduled surgical, medical, or unscheduled surgical) and presence of 3 chronic diseases (metastatic cancer, hematologic malignancy, and AIDS).

Feature Selection

After exploratory analyses, we assessed the importance of the indicators as potential predictors. First, we used feature selection methods, which are widely used to determine which predictors should be used in a model, particularly for high-dimensional data [22]. Two copies of the augmented matrix (derived from HD and PMM imputation) were fed to various feature selection methods. Our study considered in-hospital and 30-day postdischarge mortality as outcomes. Overall, we used 2 categories of supervised feature selection methods described below.

First, filter techniques evaluated the importance of a predictor by looking at data properties. Filter methods, in general, use a metric to identify irrelevant features and filter out the redundant predictors from the data matrix [57]. We selected 3 different metrics: LR beta value, relief algorithm [58], and information gain (InfGain) [59]. The relief algorithm examines the relevance of predictors based on their power to distinguish between similar patients with the same and different outcome. InfGain measures the reduction in entropy of the class variable achieved by partitioning the data based on the index predictor; relevant predictors receive a high InfGain value [60]. This ensemble of the scoring methods was then used to determine the normalized informativeness of all predictors. Aggregating these methods in one score provides a tool for comparing predictors from different aspects.

Second, we used embedded techniques to search for the optimal set of predictors. In these techniques, feature selection is embedded in the model's construction and interacts with the classifier. Least absolute shrinkage and selection operator (LASSO), used in this study, is a penalizing method in this category. LASSO regression in its objective functions considers a penalty that equals to the sum of the absolute values of the

coefficients. As absolute function (L_1 norm) is not differentiable, the estimated coefficients are close to 0, and some will be exactly 0 resulting in an automatic variable selection. For this and the next experiments, 10-fold cross-validation with 20 repeats was used (leading to 200 repetitions in total). This number of repetitions is recommended to achieve desired accuracy for prediction performance estimation [49].

Predictive Modeling

In the last assessment, we first trained group of classification models, including DT, logistic regression (LR), and random forest (RF), on the indicator and imputed data matrices and evaluate their performance for predicting desired outcomes using the area under the curve of the receiver operating characteristic (AUROC) validation metric. Thereafter, new models were trained using the augmented data matrix and their performance was compared with that of the original to determine whether the indicators have predictive power and can boost the models' predictive accuracy. We also investigated the predictive performance of SAPS-II score, and then we added indicators to these scores to examine the impact of indicators beyond SAPS-II score. It is worth mentioning that in this assessment, the absolute accuracy of the models is not of our interest, instead, the relative improvement in the performance when including indicators as input. That is, achieving the best possible mortality prediction AUROC is not the objective of this study.

Results

Population

The analyses of the first 24 hours ICU stays included 32,618 patients but decreased to 20,381 for the second 24-hour interval, as many patients were discharged after 24 hours. The third 24-hour period included 13,670 patients. Of these groups, 10.99% (3586/32,618), 13.59% (2769/20,381), and 16.19% (2213/13,670) experienced death in-hospital and 15.12% (4933/32,618), 18.26% (3722/20,381), and 21.32% (2915/13,670) experienced death within 30 days of discharge, respectively. Figure 2 demonstrates the retrospective study design.

Exploratory Analysis

Missingness rates for LTs ranges from 1.36% (445/32,618) to 88.27% (12066/13,670) in the first 72 hours after admission. Figure 3 shows the missingness rate for LTs over 3 days. Absolute basophils (BA), absolute eosinophils (EO), absolute monocytes (MO), absolute lymphocytes (LY), absolute neutrophils (NE), alanine aminotransferase (ALT), alkaline phosphatase (ALK), aspartate aminotransferase (AST), total bilirubin (TBil), and lactate (Lac) were among the less-common LTs and were missing in the profiles of more than 60% of patients.

We calculated the association between each indicator and the mortality flag. Although association values were small, on day 1, ALT, ALK, AST, and TBil stand out as the top LTs associated with both types of mortality.

Figure 2. The retrospective cohort study design. LOS: length of stay.

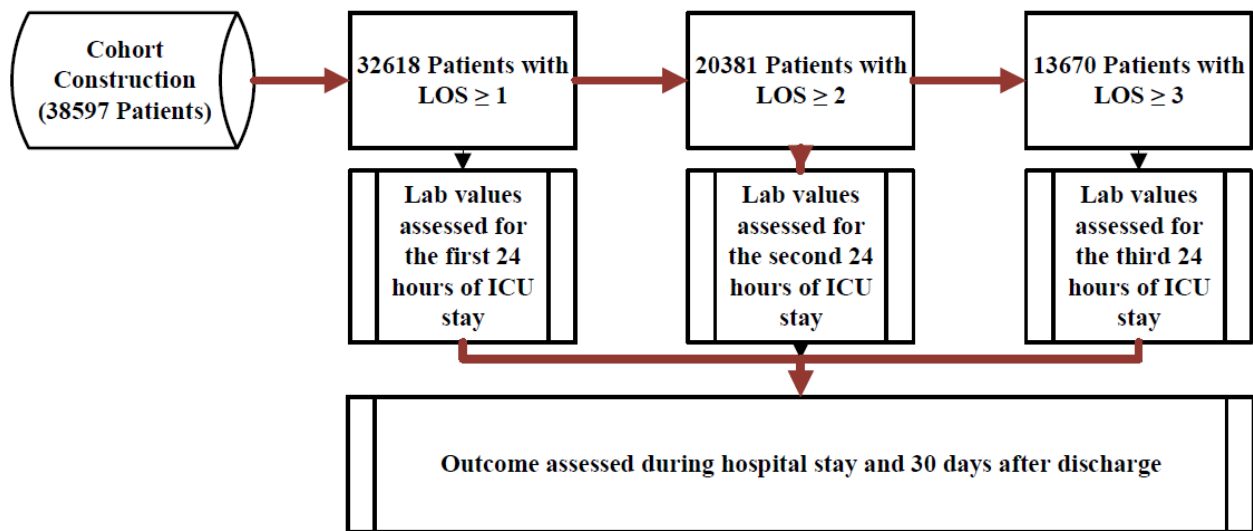
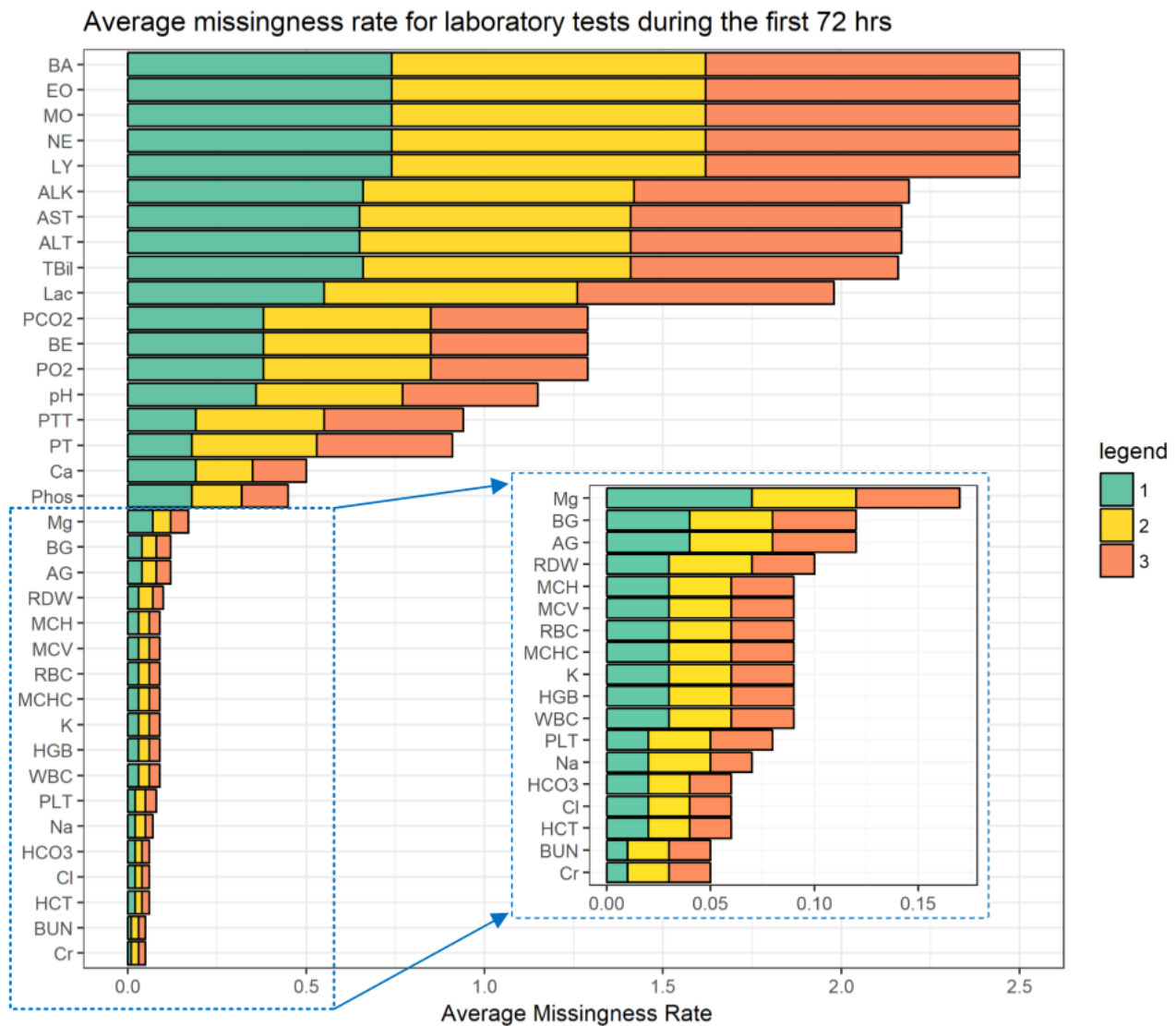


Figure 3. The average missingness rate among patients for laboratory tests in the first 72 hours of admission.



On days 2 and 3, partial pressure of carbon dioxide (PCO₂), partial pressure of oxygen (PO₂), and base excess (BE) were the top LTs associated with both mortality types. Lac also joined the top tests on day 2 for 30-day mortality. Detailed association values are provided in See [Multimedia Appendix 1](#).

Figure 4 visualizes the pairwise correlations among indicators. In total, 7 major groups of highly correlated ($\rho \geq .95$) indicators were observed in the results using Phi coefficient: (1) BA, MO, NE, EO, and LY; (2) mean corpuscular hemoglobin concentration (MCHC), red cell distribution width (RDW) mean corpuscular volume (MCV), red blood cell (RBC), and mean corpuscular hemoglobin (MCH); (3) BE, PCO₂, and PO₂; (4) TBil, ALT, AST, and ALK; (5) Blood urea nitrogen (BUN) and creatinine (Cr); (6) chloride (Cl) and bicarbonate (HCO₃); (7) partial thromboplastin time (PTT) and prothrombin time (PT).

The Spearman correlation between missingness rates and ECI was also calculated daily. Results show a statistically significant correlation between these variables (day 1: $\rho = -.233$; day 2: $\rho = -.196$; day 3: $\rho = -.184$; $P < .001$). The same assessment was

done using SAPS-II. The results were in line with the previous one and demonstrate higher correlation (day 1: $\rho = -.315$; day 2: $\rho = -.277$; day 3: $\rho = -.234$; $P < .001$). These findings are interesting as they confirm that the missingness of data is associated with patient severity of illness.

Feature Selection: Missing Data Indicators as Important Predictors

Each of the imputation methods was applied to the original dataset, and the potential informativeness of missingness indicators in comparison with actual variables was investigated using an ensemble of the most representative filter selection methods [61]: LR beta value, relief, and InfGain. Table 1 shows the top 18 variables selected on each day based on the PMM-generated imputed matrix predicting 30-day mortality. BUN, RDW, and anion gap (AG) were among the top variables in all 3 days. Indicators for TBil, phosphate (Phos), calcium (Ca), and Lac were selected on the first day, whereas indicators for Lac, BE, PO₂, and PCO₂ were among the top features on the second and third days. PTT and pH indicators were also among the important indicators on the third day.

Figure 4. Visualization of the correlation matrix for variable indicators in first 72 hours.

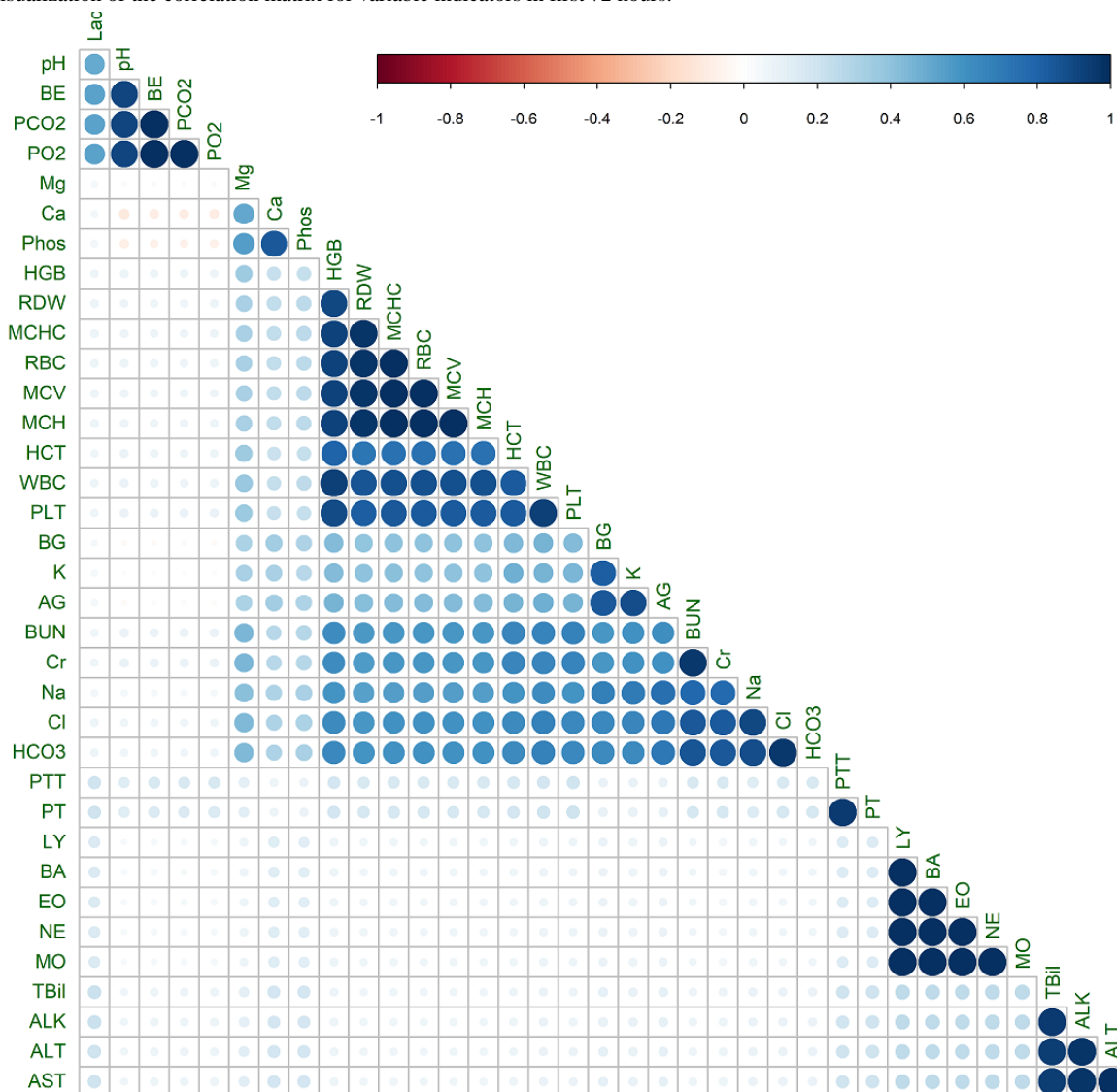


Table 1. The top 18 variables selected on each day after employing predictive mean matching imputation with regard to 30-day mortality. *I* at the beginning of the variables' names means *indicator*. Numbers represent the ranking after aggregating the ranking results from the 3 different feature selection methods.

Day 1		Day 2		Day 3	
Variable	Score	Variable	Score	Variable	Score
BUN ^a	.762397	AG ^b	.795419	RDW ^c	.748997
RDW	.680087	HCO ₃ ^d	.783337	BUN	.666667
MCHC ^e	.668965	BUN	.77677	HCO ₃	.544964
AG	.540484	BE ^f	.609532	BE	.540542
I-Ca ^g	.436429	RDW	.608711	pH	.488433
Cr ^h	.436071	I-PO ₂ ⁱ	.587151	AG	.450426
HCO ₃	.416741	I-PCO ₂	.585947	I-Lac ^j	.418716
PO ₂ ^k	.404289	I-BE	.585592	I-pH	.40463
MCV ^l	.386964	Cl ^m	.53158	Cr	.400008
I-Phos ⁿ	.374431	PT ^o	.462085	Phos	.387661
PTT ^p	.353913	Lac	.461869	I-PCO ₂	.387019
HGB ^q	.342786	Cr	.451999	I-PO ₂	.386739
pH	.32767	PTT	.424956	I-BE	.385935
Lac	.320339	Na ^r	.422474	PCO ₂	.367257
BE	.320299	Phos	.419171	NE ^s	.360791
I-Lac	.318216	I-Lac	.415475	MCV	.351266
PCO ₂	.316668	MCV	.368343	I-PTT	.338352
I-TBil ^t	.31277	MCHC	.363146	Lac	.331205

^aBUN: blood urea nitrogen.

^bAG: anion gap.

^cRDW: red cell distribution width.

^dHCO₃: bicarbonate.

^eMCHC: mean corpuscular hemoglobin concentration.

^fBE: base excess.

^gCA: calcium.

^hCr: creatinine.

ⁱPO₂: partial pressure of oxygen.

^jLac: lactate.

^kPCO₂: partial pressure of carbon dioxide.

^lMCV: mean corpuscular volume.

^mCl: chloride.

ⁿPhos: phosphate.

^oPT: prothrombin time.

^pPTT: partial thromboplastin time.

^qHGB: hemoglobin.

^rNa: sodium.

^sNE: absolute neutrophils.

^tTBil: total bilirubin.

Similar results were observed when using the HD imputation method, except that ALT and Phos were also selected on the

first and second day, respectively. Moreover, PTT and pH indicators were not among the important indicators on the third

day. Detailed results of this assessment can be found in [Multimedia Appendix 1](#).

Results for in-hospital mortality were slightly different ([Table 2](#)). Although the selected indicators were almost the same as for 30-day mortality, more indicators were selected on the first day for in-hospital mortality, implying that indicators are more associated with in-hospital mortality than 30-day mortality. Detailed results are available in [Multimedia Appendix 1](#).

To validate our previous results, we assessed the predictive power of the indicators using embedded feature selection methods. Each day, a LASSO model was trained on the augmented data from HD and PMM imputation using 10-fold cross-validation with 20 repeats. In general, the AUROC of mortality prediction (in-hospital and 30-day postdischarge) and number of selected variables decreased from days 1 to 3 ([Table 3](#)).

Moreover, prediction of in-hospital mortality resulted in higher AUROCs than 30-day mortality. Regardless of mortality type, on all days, more than 40% of the predictors selected by the best-performing model were indicators. Moreover, more than 61% of selected predictors were indicators on the third day. Sliding lambda to compromise the predictor number and model performance led to almost the same results. Generally, more than 40% of the selected predictors were indicators, and on the third day, this number increased to 61%.

Results in this section once more confirm the informativeness of missing data as missingness indicators have been selected by various feature selection methods. The high percentage of selected indicators also implies that the actual value of an LT is not always required in outcome prediction; instead, knowledge about whether the test was performed would suffice.

Predictive Modeling: Missing Data Indicators in Predictive Modeling

In the second assessment, we compared the performance of a set of 3 classification models (DT, LR, and RF) using the indicators, imputed and augmented data matrices, and SAPS-II score with or without indicators with 10-fold cross-validation over 20 repeats. We investigated whether including indicators

can improve prediction and whether indicators alone have predictive power. For our LR, the iteratively reweighted least square method was used to fit the model. The complexity parameter (CP) for DT was tuned based on the model performance. On the basis of some preliminary model fitting, we set the CP value to vary from 0 (including all variables and having a large tree) to .02 for each model and then we picked the best performance model. In all models, the best-tuned model had a CP greater than 0. [Figure 5](#) shows the AUROC with 95% CI for all 3 days with regard to 30-day mortality ([Multimedia Appendix 1](#) provides the AUROC values for 30-day mortality and in-hospital mortality).

Including indicators improved the AUROC in all modeling techniques, on average by 0.0511; the maximum improvement was 0.1209 ([Figure 5](#)). AUROC has been demonstrated as an insensitive metric, for which an increase of 0.01 suggests meaningful improvement and is clinically of interest [62-64]. Although using only indicators demonstrated reasonable performance in all scenarios (AUROC=0.6019 [0.0862]>0.5), conventional scores such as SAPS II perform better (AUROC=0.6390 [0.0853]) on their own. Therefore, models trained only on indicators are not sufficient. However, including indicators with conventional scores can improve the performance (AUROC=0.7263 [0.0578]). The SAPS-II score has information for age, heart rate, systolic blood pressure, Glasgow coma scale, temperature, mechanical ventilation administration, partial pressure of oxygen in the arterial blood (PaO₂), fraction of inspired oxygen (FiO₂), urine output, BUN, sodium (Na), potassium (K), HCO₃, TBil, white blood cells (WBCs), presence of chronic diseases, and type of admission. These results demonstrate that indicators have information beyond that included in SAPS-II.

[Figure 6](#) demonstrates the AUROC curves for LR 30-day mortality prediction on day 1.

This combination of findings provides more support for the informativeness of missing data. Employing the missing indicators in mortality prediction modeling can improve the results in comparison to not including them.

Table 2. The top 18 variables selected on each day after employing predictive mean matching imputation with regard to in-hospital mortality. *I* at the beginning of the variables names means *indicator*. Numbers represent the ranking after aggregating the ranking results from the 3 different feature selection methods.

Day 1		Day 2		Day 3	
Variable	Score	Variable	Score	Variable	Score
BUN ^a	.825715	BUN	1	RDW ^b	.75246
AG ^c	.668918	RDW	.711852	BUN	.635729
RDW	.573188	HCO ₃ ^d	.684191	BE ^e	.633926
HCO ₃	.531746	AG	.664339	HCO ₃	.62367
MCHC ^f	.507343	BE	.528778	I-BE	.595553
PCO ₂ ^g	.489483	MCHC	.503805	I-PCO ₂	.595238
Cr ^h	.480181	PT ⁱ	.453111	I-PO ₂ ^j	.594924
BE	.452599	Cl ^k	.429405	pH	.556242
I-Lac ^l	.436382	I-Lac	.425279	Phos ^m	.494694
Lac	.415773	Cr	.395266	AG	.492864
HGB ⁿ	.414263	I-PO ₂	.382404	I-pH	.470007
pH	.402466	I-PCO ₂	.381737	I-Lac	.469215
I-TBil ^o	.399363	I-BE	.381448	Cr	.415249
I-Ca	.395278	PTT ^p	.357339	Lac	.396136
I-ALT ^q	.376004	Phos	.352738	NE ^r	.338372
I-AST ^s	.375944	Na ^t	.345109	PT	.326491
LY ^u	.375163	I-PT	.333936	LY	.319146
I-ALK ^v	.366346	BG ^w	.320947	MCV ^x	.314868

^aBUN: blood urea nitrogen.

^bRDW: red cell distribution width.

^cAG: anion gap.

^dHCO₃: bicarbonate.

^eBE: base excess.

^fMCHC: mean corpuscular hemoglobin concentration.

^gPCO₂: partial pressure of carbon dioxide.

^hCr: creatinine.

ⁱPT: prothrombin time.

^jPO₂: partial pressure of oxygen.

^kCl: chloride.

^lLac: lactate.

^mPhos: phosphate.

ⁿHGB: hemoglobin.

^oTBil: total bilirubin.

^pPTT: partial prothrombin time.

^qALT: alanine transaminase.

^rNE: absolute neutrophils.

^sAST: aspartate transaminase

^tNa: sodium

^uLY: absolute lymphocytes.

^vALK: alkaline phosphatase.

^wBG: blood glucose.

^xMCV: mean corpuscular volume.

Table 3. Results from feature selection by least absolute shrinkage and selection operator (LASSO) for 3 days (area under the curve of the receiver operating characteristics are reported with the SE). The *best performing model* refers to the model with a lambda value associated with minimum cross-validation error. The adjusted model refers to a LASSO model with the largest value of lambda such that the error remains within 1 SE of the minimum.

Criteria, outcome, and imputation method	Day 1	Day 2	Day 3
AUROC^a for best performing model			
30-day mortality			
HD ^b	0.7858 (0.0033)	0.7685 (0.0041)	0.7302 (0.0043)
PMM ^c	0.7876 (0.0039)	0.7708 (0.0046)	0.7391 (0.0053)
In-hospital mortality			
HD	0.7983 (0.0040)	0.7804 (0.0046)	0.7476 (0.0042)
PMM	0.8007 (0.0047)	0.7838 (0.0049)	0.7582 (0.0054)
Indicators among selected predictors by the best performing model, n (%)			
30-day mortality			
HD	23 (43)	24 (48)	19 (707)
PMM	26 (45)	26 (47)	17 (68)
In-hospital mortality			
HD	28 (46)	29 (48)	21 (60)
PMM	29 (47)	27 (49)	24 (62)
AUROC for adjusted model			
30-day mortality			
HD	0.7826 (0.0034)	0.7646 (0.0043)	0.7262 (0.0041)
PMM	0.7840 (0.0038)	0.7667 (0.0045)	0.7339 (0.0044)
In-hospital mortality			
HD	0.7944 (0.0043)	0.7762 (0.0047)	0.7439 (0.0041)
PMM	0.7961 (0.0049)	0.7793 (0.0050)	0.7536 (0.0045)
Indicators among selected predictors by the adjusted model, n (%)			
30-day mortality			
HD	20 (45)	16 (48)	22 (67)
PMM	19 (45)	16 (52)	31 (62)
In-hospital mortality			
HD	20 (47)	13 (42)	16 (64)
PMM	18 (50)	11 (41)	16 (62)

^aAUROC: area under the curve of the receiver operating characteristic.

^bHD: hot deck.

^cPMM: predictive mean matching.

Figure 5. The 95% CIs of the area under the curve of the receiver operating characteristic for logistic regression, decision tree, and random forest models on missingness indicators, simplified acute physiology score -II, and actual variables with and without the missingness indicators.

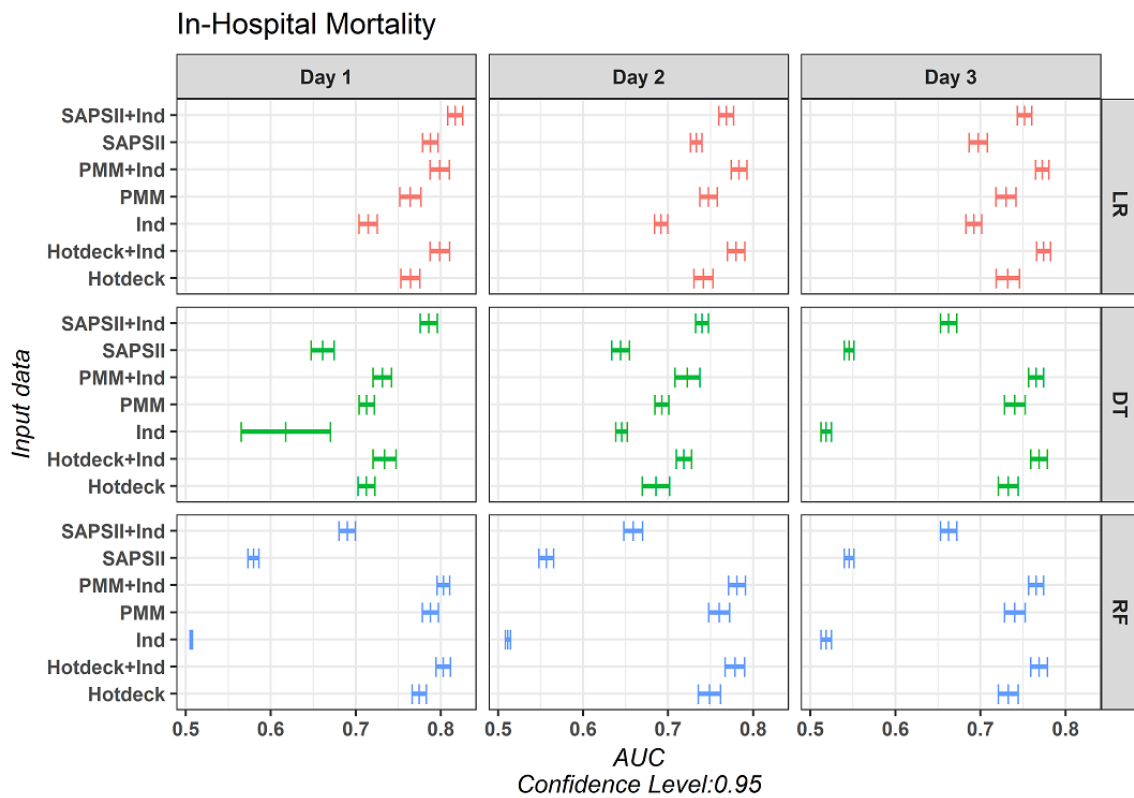
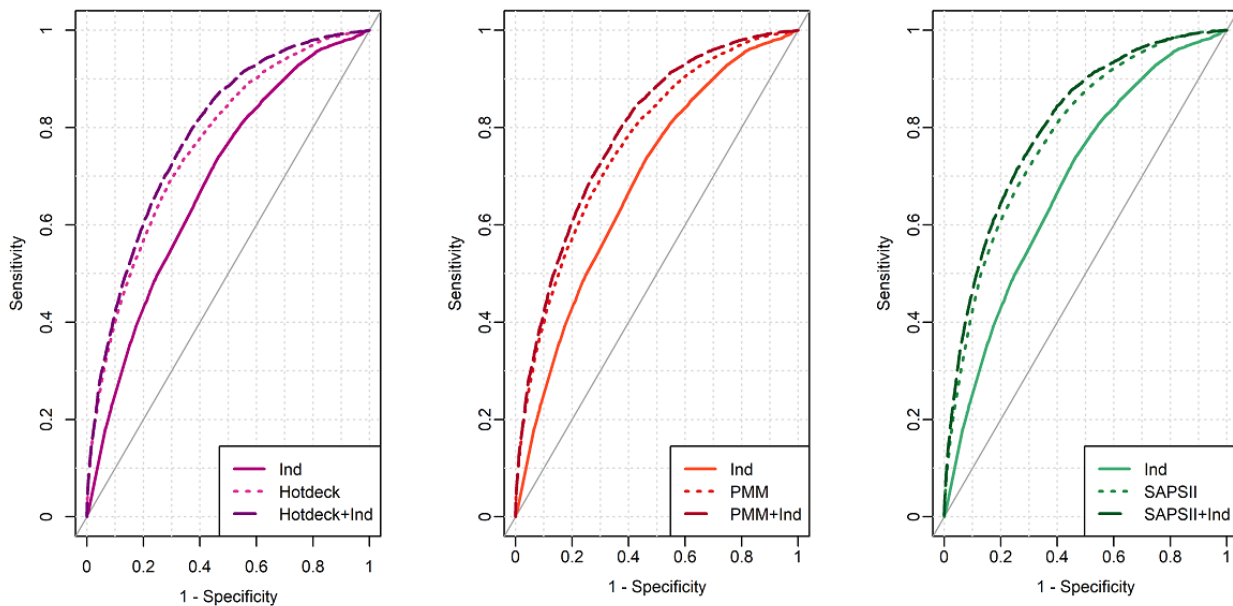


Figure 6. The receiver operating characteristic curves for logistic regression 30-day mortality prediction on day 1.



Discussion

Principal Findings

We used missingness indicators to represent missing information in patient profiles in ICU. The informativeness of these indicators was demonstrated in 3 sets of assessments. First, our exploratory analysis confirms that the missingness of data is associated with patient severity of illness or comorbidities.

Afterward, by means of feature selection methods, the predictive power of the presence of an LT in the patient profile was found to be more than the actual measured value. Finally, missingness indicators noticeably improved the performance of mortality prediction models. The high correlation observed among some of the variable indicators suggests that all the variables in a set are typically measured or ordered together. Therefore, if a patient is missing 1 variable of a set, he or she will likely be missing the others as well. This fact is well represented in all 7

groups. The first group comprises the differential WBC counts (BA, MO, NE, eosinophil; EO, and LY), which itemizes the number of basophils, monocytes, neutrophils, eosinophils, and lymphocytes among present WBCs. The second group (RDW, MCHC, MCV, RBC, and MCH) comprises tests that are used to measure the actual number of RBCs and their physical characteristics. The third group (BE, PCO₂, and PO₂) consists of blood gas components and focuses on oxygen and carbon dioxide pressure as well as excess or deficit of base levels in the blood. Tbil, ALT, AST, and ALK in the fourth group are liver enzymes [65] that are ordered when a patient is suffering from or showing symptoms of a liver-related comorbidity. BUN and Cr mainly focus on kidney function. Bicarbonate; HCO₃ and chloride; Cl are the primary measured anions in the blood. PT along with PTT are used for investigating hemostasis and are the starting points for looking into potential bleeding or clotting complications. Therefore, the presence of a clinical variable in a patient profile can represent a comorbidity in the patient. Although LTs are mainly ordered for diagnostic and prognostic reasons, studies have shown widely diverse test-ordering behavior among clinicians for similar symptoms [66-68]. Therefore, indicators could also reflect the opinions, preconceptions, and biases of the treating clinicians. In other words, by using the missingness indicators, we are learning from practice patterns rather than physiologic patterns. Therefore, indicators as introduced in this study can then be used for modeling health care process in various applications such as clinical care, clinical research, health care economics, and health care policy [21,69].

Filter methods verified the importance of some indicators with regard to our outcomes. Results also demonstrated that indicators become more and more important on ICU days 2 and 3 (Tables 1 and 2). This observation aligns with clinical practice in which ICU clinicians might try to get a complete dataset on day 1 to fully investigate the patient and understand the situation but are likely to be more selective with LT ordering on subsequent days. The Lac indicator was associated with 30-day and in-hospital mortality on the second and third day. Lactate is usually used as a biomarker for shock states. The literature has constantly reported an association between lactate levels and mortality rates among critically ill patients [70]. Our study demonstrated that just the presence of this information could represent the severity of a patient's illness, as patients with profound shock have a very high mortality rate in hospitals and ICUs [71]. Moreover, BUN [72-74], RDW [75-79], and AG [80-83] have been repeatedly determined as a risk factor of all-cause mortality and their indicators received a high score in our analysis. These results are consistent with those of Agniel et al's [21] who demonstrated that the presence of these tests have significant association with odds of 3-years survival.

The LASSO model selected indicators among the clinical predictors of in-hospital mortality and 30-day mortality, implying the predictive power of indicators. More indicators than clinical variables were selected on the third day (60%-70% of selected predictors were indicators); the assessment demonstrates that indicators from the third day are more informative than those from the first, again supporting the idea that the practice patterns diverge later during ICU stays, so there

is more variability in what gets measured. In other words, care on the first day is likely to be highly protocolized—all patients get the same tests regardless of their condition because their trajectory is still unclear. As time goes on, the patterns become more evident and ordering and prescribing practices change according to clinical need. This high percentage of selected indicators suggests that clinical variables are not always required in outcome prediction; instead, information about their presence would suffice.

The last assessment demonstrated that models trained on indicators alone in some scenarios have reasonable performance (for instance, in 30-day mortality prediction with LR, we achieved AUROC of 0.6836 [0.012]). These results imply that by considering missing data as noise or a random artifact, we can lose valuable information about patient outcomes. Moreover, indicators improved the AUROCs in most scenarios. Researchers in this field are looking for predictors that can be included in the models to improve the prediction results. Having a low-dimensional set of typical predictors plus these missing data indicators can actually lead to performance comparable with that achieved using typical predictors plus other potentially useful predictors identified a priori by medical researchers: First, in comparison with including extra numeric predictors, the computational load for performing mathematical calculations on binary values such as indicators is usually less. Second, binary data require less computational memory than numbers when performing data mining techniques. Finally, for some important clinical variables, storing the missing data indicators instead of the actual value better protects patient privacy while preserving predictive power. In other words, less privacy concern is expected in a situation when the type of test is disclosed rather than the actual test result. The comparative analyses on the predictive models showed that missing data indicators could improve the prediction models' performance. Although literature considers a small increase (0.01) in AUROC meaningful and of clinical interest (because of insensitivity of AUROC) [62,64], including the indicators in our study could improve the average AUROC by 0.0511. Thus, missing data indicators can be introduced as informative predictors and be used to learn from. In other words, these indicators can be representative of physicians' and patients' opinions during the health care process. Furthermore, the overall model performance decreased over time perhaps implying that patients' data on the first 24-hour has the highest level of information. The same pattern was also observed in the previous assessment. According to these observations, we can infer that presence or absence of a variable can be used in predicting patients' severity of illness.

Strengths and Limitations of the Project

A significant strength of this study is its new insight on missing data in a real-world ICU database. The results confirm the predictive power of some indicators and their advantage over actual values in predictive modeling. The findings further clarify the factors associated with lack of data collection such as the healthier status of a patient or practice patterns of clinicians. These insights, in turn, can be used to design models that consider missing data and benefit from the hidden information. On the basis of our results, missingness indicators can be introduced as potential predictors of ICU patients' outcome.

Despite the strength, significance, and novel nature of this study, there also exist limitations that cannot be overlooked. First, because of the nature of ICUs, the amount of missing data in MIMIC is less than that from a general ward. Therefore, our study may not fully demonstrate the informativeness of these indicators. Moreover, adding the indicators of interest to the actual data matrix increases the dimension of the matrix and may become computationally burdensome. Using other imputation methods, the power of missing data indicators may vary but this was beyond the scope of our study, which focused on providing evidence on missing data informativeness.

Perspectives for Future Work

Although our study demonstrates that missingness indicators are informative and have predictive power in mortality prediction in ICU, further studies are required to investigate their power in predicting other clinical outcomes. Future researchers can investigate the association between missingness patterns and patient diagnosis. They can also consider more sensitive criteria such as net reclassification or integrated discrimination improvements while preserving improvement in the AUROC as the first criterion. Moreover, as this study looked at the 3 days in the ICU independently, one can

investigate if the missing data on a particular day are still informative given all the clinical and indicator variables from previous days. These future studies should also investigate the effect of missing rate on the predictive power of indicators. Another area of future work is examining the test-ordering behavior among clinicians, by using missingness indicators.

Conclusions

Our study has demonstrated that the missingness of data itself might be informative in ICU and might have added predictive value beyond observed data alone. Moreover, indicators for variables with higher missingness rates had more predictive power. In practice, the lack of a set of symptoms might lead health professionals to conclude that a particular set of tests is not required at the current stage. Therefore, these missing data are not a random occurrence. This study showed that the number of comorbidities is associated with a decreased rate of missing data. Therefore, rudimentary treatments of missing data (eg, CCA) can cause bias toward sicker patients. The study is also notable because it provided new insight about the informativeness of missing data and described how this information could be used in predicting mortality.

Acknowledgments

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2014-04743, RGPIN-2014-05911) and Early Researcher Award (Ministry of Research and Innovation, Government of Ontario).

Authors' Contributions

Study conception and design were conducted by AS, JAD, DMM, and JL. AS extracted data and performed the data analysis. Interpretation of the results was provided by all authors. All authors contributed in writing the paper and approved the final version of the review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed results.

[[PDF File \(Adobe PDF File\), 165 KB - medinform_v7i1e11605_app1.pdf](#)]

References

1. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc* 2013 Jan 1;20(1):144-151 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2011-000681](https://doi.org/10.1136/amiajnl-2011-000681)] [Medline: [22733976](https://pubmed.ncbi.nlm.nih.gov/22733976/)]
2. Chan KS, Fowles JB, Weiner JP. Review: electronic health records and the reliability and validity of quality measures: a review of the literature. *Med Care Res Rev* 2010 Oct;67(5):503-527. [doi: [10.1177/1077558709359007](https://doi.org/10.1177/1077558709359007)] [Medline: [20150441](https://pubmed.ncbi.nlm.nih.gov/20150441/)]
3. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013 Oct;46(5):830-836 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2013.06.010](https://doi.org/10.1016/j.jbi.2013.06.010)] [Medline: [23820016](https://pubmed.ncbi.nlm.nih.gov/23820016/)]
4. Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record derived data. *EGEMS (Wash DC)* 2013;1(3):1035 [[FREE Full text](#)] [doi: [10.13063/2327-9214.1035](https://doi.org/10.13063/2327-9214.1035)] [Medline: [25848578](https://pubmed.ncbi.nlm.nih.gov/25848578/)]
5. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012 Oct 4;367(14):1355-1360 [[FREE Full text](#)] [doi: [10.1056/NEJMSr1203730](https://doi.org/10.1056/NEJMSr1203730)] [Medline: [23034025](https://pubmed.ncbi.nlm.nih.gov/23034025/)]

6. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Br Med J* 2009;338:b2393 [FREE Full text] [Medline: 19564179]
7. Haukoos JS, Newgard CD. Advanced statistics: missing data in clinical research--part 1: an introduction and conceptual framework. *Acad Emerg Med* 2007 Jul;14(7):662-668 [FREE Full text] [doi: 10.1197/j.aem.2006.11.037] [Medline: 17538078]
8. Newgard CD, Haukoos JS. Advanced statistics: missing data in clinical research--part 2: multiple imputation. *Acad Emerg Med* 2007 Jul;14(7):669-678 [FREE Full text] [doi: 10.1197/j.aem.2006.11.038] [Medline: 17595237]
9. Pringle M, Ward P, Chilvers C. Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer. *Br J Gen Pract* 1995 Oct;45(399):537-541 [FREE Full text] [Medline: 7492423]
10. Thiru K, de Lusignan S, Hague N. Have the completeness and accuracy of computer medical records in general practice improved in the last five years? The report of a two-practice pilot study. *Health Informatics J* 2016 Jul 25;5(4):224-232. [doi: 10.1177/146045829900500410]
11. Forster M, Bailey C, Brinkhof MW, Graber C, Boulle A, Spohr M, ART-LINC collaboration of International Epidemiological Databases to Evaluate AIDS. Electronic medical record systems, data quality and loss to follow-up: survey of antiretroviral therapy programmes in resource-limited settings. *Bull World Health Organ* 2008 Dec;86(12):939-947 [FREE Full text] [Medline: 19142294]
12. Jones RB, Hedley AJ. A computer in the diabetic clinic. Completeness of data in a clinical information system for diabetes. *Pract Diab Int* 1986 Nov;3(6):295-296. [doi: 10.1002/pdi.1960030610]
13. Porcheret M, Hughes R, Evans D, Jordan K, Whitehurst T, Ogden H, North Staffordshire General Practice Research Network. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. *J Am Med Inform Assoc* 2004;11(1):78-86 [FREE Full text] [doi: 10.1197/jamia.M1362] [Medline: 14527973]
14. Soto CM, Kleinman KP, Simon SR. Quality and correlates of medical record documentation in the ambulatory care setting. *BMC Health Serv Res* 2002 Dec 10;2(1):22 [FREE Full text] [Medline: 12473161]
15. Tang PC, LaRosa MP, Gorden SM. Use of computer-based records, completeness of documentation, and appropriateness of documented clinical decisions. *J Am Med Inform Assoc* 1999 May 01;6(3):245-251. [doi: 10.1136/jamia.1999.0060245]
16. Jensen RE, Chan KS, Weiner JP, Fowles JB, Neale SM. Implementing electronic health record-based quality measures for developmental screening. *Pediatrics* 2009 Oct;124(4):e648-e654. [doi: 10.1542/peds.2008-3091] [Medline: 19786425]
17. Botsis T, Hartvigsen G, Chen F, Weng C. Secondary use of EHR: data quality issues and informatics opportunities. *AMIA Jt Summits Transl Sci Proc* 2010 Mar 01;2010:1-5 [FREE Full text] [Medline: 21347133]
18. Sharafoddini A, Dubin JA, Lee J. Patient similarity in prediction models based on health data: a scoping review. *JMIR Med Inform* 2017 Mar 03;5(1):e7 [FREE Full text] [doi: 10.2196/medinform.6730] [Medline: 28258046]
19. Rusanov A, Weiskopf NG, Wang S, Weng C. Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research. *BMC Med Inform Decis Mak* 2014 Jun 11;14:51. [doi: 10.1186/1472-6947-14-51] [Medline: 24916006]
20. Weiskopf NG, Rusanov A, Weng C. Sick patients have more data: the non-random completeness of electronic health records. *AMIA Annu Symp Proc* 2013;2013:1472-1477 [FREE Full text] [Medline: 24551421]
21. Agniel D, Kohane I, Weber G. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. *Br Med J* 2018 Dec 30;361:k1479 [FREE Full text] [doi: 10.1136/bmj.k1479] [Medline: 29712648]
22. Kuhn M, Johnson K. *Applied Predictive Modeling*. New York: Springer; 2013.
23. Lee J, Maslove DM. Using information theory to identify redundancy in common laboratory tests in the intensive care unit. *BMC Med Inform Decis Mak* 2015 Jul 31;15:59 [FREE Full text] [doi: 10.1186/s12911-015-0187-x] [Medline: 26227625]
24. Oliveira AM, Oliveira MV, Souza CL. Prevalence of unnecessary laboratory tests and related avoidable costs in intensive care unit. *J Bras Patol Med Lab* 2014;50:410-416. [doi: 10.5935/1676-2444.20140049]
25. Cismondi F, Celi LA, Fialho AS, Vieira SM, Reti SR, Sousa JM, et al. Reducing unnecessary lab testing in the ICU with artificial intelligence. *Int J Med Inform* 2013 May;82(5):345-358 [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.11.017] [Medline: 23273628]
26. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365(9453):82-93. [doi: 10.1016/S0140-6736(04)17670-8] [Medline: 15639683]
27. Doyle JD, Garmon EH. American Society of Anesthesiologists Classification (ASA Class). *StatPearls* 2018. [Medline: 28722969]
28. Vesin A, Azoulay E, Ruckly S, Vignoud L, Rusinová K, Benoit D, et al. Reporting and handling missing values in clinical studies in intensive care units. *Intensive Care Med* 2013 Aug;39(8):1396-1404. [doi: 10.1007/s00134-013-2949-1] [Medline: 23685609]
29. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. Hoboken, New Jersey: Wiley; 2011.
30. Baraldi AN, Enders CK. An introduction to modern missing data analyses. *J Sch Psychol* 2010 Feb;48(1):5-37. [doi: 10.1016/j.jsp.2009.10.001] [Medline: 20006986]
31. Stockdale M, Royal K. Missing data as a validity threat for medical and healthcare education research: problems and solutions. *Int J Health Care* 2016 Jun 23;2(2). [doi: 10.5430/ijh.v2n2p67]

32. Myers TA. Goodbye, listwise deletion: presenting hot deck imputation as an easy and effective tool for handling missing data. *Communication Methods and Measures* 2011 Oct;5(4):297-310. [doi: [10.1080/19312458.2011.624490](https://doi.org/10.1080/19312458.2011.624490)]
33. Pigott TD. A review of methods for missing data. *Educ Res Eval* 2001 Dec 1;7(4):353-383. [doi: [10.1076/edre.7.4.353.8937](https://doi.org/10.1076/edre.7.4.353.8937)]
34. Roth PL. Missing data - a conceptual review for applied psychologists. *Pers Psychol* 1994;47(3):537-560. [doi: [10.1111/j.1744-6570.1994.tb01736.x](https://doi.org/10.1111/j.1744-6570.1994.tb01736.x)]
35. Gorelick MH. Bias arising from missing data in predictive models. *J Clin Epidemiol* 2006 Oct;59(10):1115-1123. [doi: [10.1016/j.jclinepi.2004.11.029](https://doi.org/10.1016/j.jclinepi.2004.11.029)] [Medline: [16980153](https://pubmed.ncbi.nlm.nih.gov/16980153/)]
36. Schulman CS, Staul L. Standards for frequency of measurement and documentation of vital signs and physical assessments. *Crit Care Nurse* 2010 Jun;30(3):74-76 [FREE Full text] [doi: [10.4037/ccn2010406](https://doi.org/10.4037/ccn2010406)] [Medline: [20515885](https://pubmed.ncbi.nlm.nih.gov/20515885/)]
37. Asiimwe SB, Okello S, Moore CC. Frequency of vital signs monitoring and its association with mortality among adults with severe sepsis admitted to a general medical ward in Uganda. *PLoS One* 2014;9(2):e89879 [FREE Full text] [doi: [10.1371/journal.pone.0089879](https://doi.org/10.1371/journal.pone.0089879)] [Medline: [24587094](https://pubmed.ncbi.nlm.nih.gov/24587094/)]
38. Cardona-Morrell M, Nicholson M, Hillman K. Vital Signs: From Monitoring to Prevention of Deterioration in General Wards. In: Vincent JL, editor. *Annual Update In Intensive Care And Emergency Medicine 2015*. Cham: Springer; 2018.
39. Frassica JJ. Frequency of laboratory test utilization in the intensive care unit and its implications for large scale data collection efforts. *AMIA Annu Symp Proc* 2003:844 [FREE Full text] [Medline: [14728349](https://pubmed.ncbi.nlm.nih.gov/14728349/)]
40. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035 [FREE Full text] [doi: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35)] [Medline: [27219127](https://pubmed.ncbi.nlm.nih.gov/27219127/)]
41. Harrell M, Fabbri D, Levy M. Evaluating EHR Data Availability for Cohort Selection in Retrospective Studies. 2016 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); 2016; Chicago, IL, USA p. 4-7. [doi: [10.1109/ICHI.2016.68](https://doi.org/10.1109/ICHI.2016.68)]
42. Little RJ. Missing-data adjustments in large surveys. *J Bus Econ Stat* 1988 Jul;6(3):287-296. [doi: [10.1080/07350015.1988.10509663](https://doi.org/10.1080/07350015.1988.10509663)]
43. Zhou XH, Eckert GJ, Tierney WM. Multiple imputation in public health research. *Stat Med* 2001;20(9-10):1541-1549. [Medline: [11343373](https://pubmed.ncbi.nlm.nih.gov/11343373/)]
44. Buuren SV. *Flexible Imputation Of Missing Data*. New York: Chapman and Hall/CRC; 2018.
45. Andridge RR, Little RJ. A review of hot deck imputation for survey non-response. *Int Stat Rev* 2010 Apr;78(1):40-64 [FREE Full text] [doi: [10.1111/j.1751-5823.2010.00103.x](https://doi.org/10.1111/j.1751-5823.2010.00103.x)] [Medline: [21743766](https://pubmed.ncbi.nlm.nih.gov/21743766/)]
46. Abraham W, Russell D. Missing data: a review of current methods and applications in epidemiological research. *Curr Opin Psychiatr* 2004 Jul;17(4):315-321. [doi: [10.1097/01.yco.0000133836.34543.7e](https://doi.org/10.1097/01.yco.0000133836.34543.7e)]
47. Groenwold RH, White IR, Donders AR, Carpenter JR, Altman DG, Moons KG. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. *Can Med Assoc J* 2012 Aug 07;184(11):1265-1269 [FREE Full text] [doi: [10.1503/cmaj.110977](https://doi.org/10.1503/cmaj.110977)] [Medline: [22371511](https://pubmed.ncbi.nlm.nih.gov/22371511/)]
48. Knol MJ, Janssen KJ, Donders AR, Egberts AC, Heerdink ER, Grobbee DE, et al. Unpredictable bias when using the missing indicator method or complete case analysis for missing confounder values: an empirical example. *J Clin Epidemiol* 2010 Jul;63(7):728-736. [doi: [10.1016/j.jclinepi.2009.08.028](https://doi.org/10.1016/j.jclinepi.2009.08.028)] [Medline: [20346625](https://pubmed.ncbi.nlm.nih.gov/20346625/)]
49. Steyerberg E. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer-Verlag; 2008.
50. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998 Jan;36(1):8-27. [Medline: [9431328](https://pubmed.ncbi.nlm.nih.gov/9431328/)]
51. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-383. [Medline: [3558716](https://pubmed.ncbi.nlm.nih.gov/3558716/)]
52. Menendez ME, Neuhaus V, van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. *Clin Orthop Relat Res* 2014 Sep;472(9):2878-2886 [FREE Full text] [doi: [10.1007/s11999-014-3686-7](https://doi.org/10.1007/s11999-014-3686-7)] [Medline: [24867450](https://pubmed.ncbi.nlm.nih.gov/24867450/)]
53. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. *Med Care* 2004 Apr;42(4):355-360. [Medline: [15076812](https://pubmed.ncbi.nlm.nih.gov/15076812/)]
54. Farley JF, Harley CR, Devine JW. A comparison of comorbidity measurements to predict healthcare expenditures. *Am J Manag Care* 2006 Feb;12(2):110-119 [FREE Full text] [Medline: [16464140](https://pubmed.ncbi.nlm.nih.gov/16464140/)]
55. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012 Dec;50(12):1109-1118. [doi: [10.1097/MLR.0b013e31825f64d0](https://doi.org/10.1097/MLR.0b013e31825f64d0)] [Medline: [22929993](https://pubmed.ncbi.nlm.nih.gov/22929993/)]
56. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *J Am Med Assoc* 1993;270(24):2957-2963. [Medline: [8254858](https://pubmed.ncbi.nlm.nih.gov/8254858/)]
57. Saes Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007 Oct 1;23(19):2507-2517 [FREE Full text] [doi: [10.1093/bioinformatics/btm344](https://doi.org/10.1093/bioinformatics/btm344)] [Medline: [17720704](https://pubmed.ncbi.nlm.nih.gov/17720704/)]
58. Robnik-Sikonja M, Kononenko I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach Learn* 2003;53(1-2):23-69. [doi: [10.1023/A:1025667309714](https://doi.org/10.1023/A:1025667309714)]

59. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 2005 Aug;27(8):1226-1238. [doi: [10.1109/TPAMI.2005.159](https://doi.org/10.1109/TPAMI.2005.159)] [Medline: [16119262](https://pubmed.ncbi.nlm.nih.gov/16119262/)]
60. Mitchell T. *Machine Learning*. New York: McGraw-Hill Education; 1997.
61. Aggarwal CC. *Data Classification: Algorithms And Applications*. New York: Chapman and Hall/CRC; 2018.
62. Martens FK, Tonk EC, Kers JG, Janssens AC. Small improvement in the area under the receiver operating characteristic curve indicated small changes in predicted risks. *J Clin Epidemiol* 2016 Nov;79:159-164. [doi: [10.1016/j.jclinepi.2016.07.002](https://doi.org/10.1016/j.jclinepi.2016.07.002)] [Medline: [27430730](https://pubmed.ncbi.nlm.nih.gov/27430730/)]
63. Cook NR. Response to letters regarding article, "use and misuse of the receiver operating characteristic curve in risk prediction". *Circulation* 2007 Aug 07;116(6):e134-e134. [doi: [10.1161/Circulationaha.107.715359](https://doi.org/10.1161/Circulationaha.107.715359)]
64. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012 Sep 15;176(6):473-481 [FREE Full text] [doi: [10.1093/aje/kws207](https://doi.org/10.1093/aje/kws207)] [Medline: [22875755](https://pubmed.ncbi.nlm.nih.gov/22875755/)]
65. Gowda S, Desai PB, Hull VV, Math AA, Vernekar SN, Kulkarni SS. A review on laboratory liver function tests. *Pan Afr Med J* 2009 Nov 22;3:17 [FREE Full text] [Medline: [21532726](https://pubmed.ncbi.nlm.nih.gov/21532726/)]
66. Wennberg JE. Dealing with medical practice variations: a proposal for action. *Health Aff (Millwood)* 1984;3(2):6-32. [Medline: [6432667](https://pubmed.ncbi.nlm.nih.gov/6432667/)]
67. Daniels M, Schroeder SA. Variation among Physicians in Use of Laboratory Tests II. Relation to Clinical Productivity and Outcomes of Care. *Medical Care* 1977;15(6):482-487. [doi: [10.1097/00005650-197706000-00004](https://doi.org/10.1097/00005650-197706000-00004)]
68. Solomon DH, Hashimoto H, Daltroy L, Liang MH. Techniques to improve physicians' use of diagnostic tests: a new conceptual framework. *J Am Med Assoc* 1998 Dec 16;280(23):2020-2027. [Medline: [9863854](https://pubmed.ncbi.nlm.nih.gov/9863854/)]
69. Sharafoddini A, Dubin JA, Lee J. Finding Similar Patient Subpopulations in the ICU Using Laboratory Test Ordering Patterns. In: *Proceedings of the 2018 7th International Conference on Bioinformatics and Biomedical Science*. 2018 Presented at: 2018 7th International Conference on Bioinformatics and Biomedical Science; 2018; Shenzhen, China. [doi: [10.1145/3239264.3239277](https://doi.org/10.1145/3239264.3239277)]
70. Zhang Z, Xu X. Lactate clearance is a useful biomarker for the prediction of all-cause mortality in critically ill patients: a systematic review and meta-analysis*. *Crit Care Med* 2014 Sep;42(9):2118-2125. [doi: [10.1097/CCM.0000000000000405](https://doi.org/10.1097/CCM.0000000000000405)] [Medline: [24797375](https://pubmed.ncbi.nlm.nih.gov/24797375/)]
71. Levinson AT, Casserly BP, Levy MM. Reducing mortality in severe sepsis and septic shock. *Semin Respir Crit Care Med* 2011 Apr;32(2):195-205. [doi: [10.1055/s-0031-1275532](https://doi.org/10.1055/s-0031-1275532)] [Medline: [21506056](https://pubmed.ncbi.nlm.nih.gov/21506056/)]
72. Beier K, Eppanapally S, Bazick HS, Chang D, Mahadevappa K, Gibbons FK, et al. Elevation of blood urea nitrogen is predictive of long-term mortality in critically ill patients independent of "normal" creatinine. *Crit Care Med* 2011 Feb;39(2):305-313 [FREE Full text] [doi: [10.1097/CCM.0b013e3181ffe22a](https://doi.org/10.1097/CCM.0b013e3181ffe22a)] [Medline: [21099426](https://pubmed.ncbi.nlm.nih.gov/21099426/)]
73. Cauthen CA, Lipinski MJ, Abbate A, Appleton D, Nusca A, Varma A, et al. Relation of blood urea nitrogen to long-term mortality in patients with heart failure. *Am J Cardiol* 2008 Jun 01;101(11):1643-1647. [doi: [10.1016/j.amjcard.2008.01.047](https://doi.org/10.1016/j.amjcard.2008.01.047)] [Medline: [18489944](https://pubmed.ncbi.nlm.nih.gov/18489944/)]
74. Kajimoto K, Sato N, Takano T, Investigators of the Acute Decompensated Heart Failure Syndromes (ATTEND) registry. Relation between elevated blood urea nitrogen, clinical features or comorbidities, and clinical outcome in patients hospitalized for acute heart failure syndromes. *Int J Cardiol* 2015 Dec 15;201:311-314. [doi: [10.1016/j.ijcard.2015.08.061](https://doi.org/10.1016/j.ijcard.2015.08.061)] [Medline: [26301667](https://pubmed.ncbi.nlm.nih.gov/26301667/)]
75. Bazick HS, Chang D, Mahadevappa K, Gibbons FK, Christopher KB. Red cell distribution width and all-cause mortality in critically ill patients. *Crit Care Med* 2011 Aug;39(8):1913-1921 [FREE Full text] [doi: [10.1097/CCM.0b013e31821b85c6](https://doi.org/10.1097/CCM.0b013e31821b85c6)] [Medline: [21532476](https://pubmed.ncbi.nlm.nih.gov/21532476/)]
76. Hunziker S, Celi LA, Lee J, Howell MD. Red cell distribution width improves the simplified acute physiology score for risk prediction in unselected critically ill patients. *Crit Care* 2012 May 18;16(3):R89 [FREE Full text] [doi: [10.1186/cc11351](https://doi.org/10.1186/cc11351)] [Medline: [22607685](https://pubmed.ncbi.nlm.nih.gov/22607685/)]
77. Patel KV, Semba RD, Ferrucci L, Newman AB, Fried LP, Wallace RB, et al. Red cell distribution width and mortality in older adults: a meta-analysis. *J Gerontol A Biol Sci Med Sci* 2010 Mar;65(3):258-265 [FREE Full text] [doi: [10.1093/gerona/glp163](https://doi.org/10.1093/gerona/glp163)] [Medline: [19880817](https://pubmed.ncbi.nlm.nih.gov/19880817/)]
78. Purtle SW, Moromizato T, McKane CK, Gibbons FK, Christopher KB. The association of red cell distribution width at hospital discharge and out-of-hospital mortality following critical illness*. *Crit Care Med* 2014 Apr;42(4):918-929. [doi: [10.1097/CCM.0000000000000118](https://doi.org/10.1097/CCM.0000000000000118)] [Medline: [24448196](https://pubmed.ncbi.nlm.nih.gov/24448196/)]
79. Şenol K, Saylam B, Kocaay F, Tez M. Red cell distribution width as a predictor of mortality in acute pancreatitis. *Am J Emerg Med* 2013 Apr;31(4):687-689. [doi: [10.1016/j.ajem.2012.12.015](https://doi.org/10.1016/j.ajem.2012.12.015)] [Medline: [23399348](https://pubmed.ncbi.nlm.nih.gov/23399348/)]
80. Ahn SY, Ryu J, Baek SH, Han JW, Lee JH, Ahn S, et al. Serum anion gap is predictive of mortality in an elderly population. *Exp Gerontol* 2014 Feb;50:122-127. [doi: [10.1016/j.exger.2013.12.002](https://doi.org/10.1016/j.exger.2013.12.002)] [Medline: [24333141](https://pubmed.ncbi.nlm.nih.gov/24333141/)]
81. Kim MJ, Kim YH, Sol IS, Kim SY, Kim JD, Kim HY, et al. Serum anion gap at admission as a predictor of mortality in the pediatric intensive care unit. *Sci Rep* 2017 May 03;7(1):1456 [FREE Full text] [doi: [10.1038/s41598-017-01681-9](https://doi.org/10.1038/s41598-017-01681-9)] [Medline: [28469150](https://pubmed.ncbi.nlm.nih.gov/28469150/)]

82. Lee SW, Kim S, Na KY, Cha R, Kang SW, Park CW, et al. Serum anion gap predicts all-cause mortality in patients with advanced chronic kidney disease: a retrospective analysis of a randomized controlled study. *PLoS One* 2016;11(6):e0156381 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0156381](https://doi.org/10.1371/journal.pone.0156381)] [Medline: [27249416](#)]
83. Sahu A, Cooper HA, Panza JA. The initial anion gap is a predictor of mortality in acute myocardial infarction. *Coron Artery Dis* 2006 Aug;17(5):409-412. [Medline: [16845247](#)]

Abbreviations

ACA: available case analysis
AG: anion gap
ALK: alkaline phosphatase
ALT: alanine aminotransferase
ASA: anesthesiologists physical status
AST: aspartate aminotransferase
AUROC: area under the curve of the receiver operating characteristic
BA: basophils
BE: base excess
BG: blood glucose
BUN: blood urea nitrogen
Ca: calcium
CCA: complete case analysis
Cl: chloride
CP: complexity parameter
Cr: creatinine
DT: decision tree
ECI: Elixhauser Comorbidity Index
EHR: electronic health record
EO: eosinophils
FiO₂: fraction of inspired oxygen
HCO₃: bicarbonate
HCT: hematocrit
HD: hot deck
HGB: hemoglobin
ICU: intensive care unit
InfGain: information gain
K: potassium
Lac: lactate
LASSO: least absolute shrinkage and selection operator
LR: logistic regression
LT: laboratory test
LY: lymphocytes
MAR: missing at random
MCAR: missing completely at random
MCH: mean corpuscular hemoglobin
MCHC: mean corpuscular hemoglobin concentration
MCV: mean corpuscular volume
Mg: magnesium
MIMIC: Medical Information Mart for Intensive Care
ML: machine learning
MO: monocytes
Na: sodium
NE: neutrophils
PaO₂: partial pressure of oxygen in the arterial blood
PCO₂: partial pressure of carbon dioxide
Phos: phosphate
PLT: platelet count
PMM: predictive mean matching
PO₂: partial pressure of oxygen
PT: prothrombin time

PTT: partial thromboplastin time
RBC: red blood cell
RDW: red cell distribution width
RF: random forest
SAPS-II: Simplified Acute Physiology Score II
TBil: total bilirubin
WBC: white blood cell

Edited by G Eysenbach; submitted 17.07.18; peer-reviewed by Z Zhang, K Tingay; comments to author 08.10.18; revised version received 30.10.18; accepted 30.10.18; published 08.01.19.

Please cite as:

Sharafoddini A, Dubin JA, Maslove DM, Lee J

A New Insight Into Missing Data in Intensive Care Unit Patient Profiles: Observational Study

JMIR Med Inform 2019;7(1):e11605

URL: <http://medinform.jmir.org/2019/1/e11605/>

doi: [10.2196/11605](https://doi.org/10.2196/11605)

PMID: [30622091](https://pubmed.ncbi.nlm.nih.gov/30622091/)

©Anis Sharafoddini, Joel A Dubin, David M Maslove, Joon Lee. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 08.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Connected Intensive Care Unit Patient: Exploratory Analyses and Cohort Discovery From a Critical Care Telemedicine Database

Patrick Essay¹, MS; Tala B Shahin², BS; Baran Balkan¹, BS; Jarrod Mosier^{3,4}, MD, FACEP, FCCM; Vignesh Subbian^{5,6}, PhD

¹College of Engineering, The University of Arizona, Tucson, AZ, United States

²College of Medicine - Tucson, The University of Arizona, Tucson, AZ, United States

³Division of Pulmonary, Allergy, Critical Care, and Sleep, Department of Medicine, The University of Arizona, Tucson, AZ, United States

⁴Department of Emergency Medicine, The University of Arizona, Tucson, AZ, United States

⁵Department of Systems and Industrial Engineering, The University of Arizona, Tucson, AZ, United States

⁶Department of Biomedical Engineering, The University of Arizona, Tucson, AZ, United States

Corresponding Author:

Patrick Essay, MS

College of Engineering

The University of Arizona

1127 East James E Rogers Way

Tucson, AZ, 85721-0020

United States

Phone: 1 4024305524

Email: pessay@email.arizona.edu

Abstract

Background: Many intensive care units (ICUs) utilize telemedicine in response to an expanding critical care patient population, off-hours coverage, and intensivist shortages, particularly in rural facilities. Advances in digital health technologies, among other reasons, have led to the integration of active, well-networked critical care telemedicine (tele-ICU) systems across the United States, which in turn, provide the ability to generate large-scale remote monitoring data from critically ill patients.

Objective: The objective of this study was to explore opportunities and challenges of utilizing multisite, multimodal data acquired through critical care telemedicine. Using a publicly available tele-ICU, or electronic ICU (eICU), database, we illustrated the quality and potential uses of remote monitoring data, including cohort discovery for secondary research.

Methods: Exploratory analyses were performed on the eICU Collaborative Research Database that includes deidentified clinical data collected from adult patients admitted to ICUs between 2014 and 2015. Patient and ICU characteristics, top admission diagnoses, and predictions from clinical scoring systems were extracted and analyzed. Additionally, a case study on respiratory failure patients was conducted to demonstrate research prospects using tele-ICU data.

Results: The eICU database spans more than 200 hospitals and over 139,000 ICU patients across the United States with wide-ranging clinical data and diagnoses. Although mixed medical-surgical ICU was the most common critical care setting, patients with cardiovascular conditions accounted for more than 20% of ICU stays, and those with neurological or respiratory illness accounted for nearly 15% of ICU unit stays. The case study on respiratory failure patients showed that cohort discovery using the eICU database can be highly specific, albeit potentially limiting in terms of data provenance and sparsity for certain types of clinical questions.

Conclusions: Large-scale remote monitoring data sources, such as the eICU database, have a strong potential to advance the role of critical care telemedicine by serving as a testbed for secondary research as well as for developing and testing tools, including predictive and prescriptive analytical solutions and decision support systems. The resulting tools will also inform coordination of care for critically ill patients, intensivist coverage, and the overall process of critical care telemedicine.

(*JMIR Med Inform* 2019;7(1):e13006) doi:[10.2196/13006](https://doi.org/10.2196/13006)

KEYWORDS

telemedicine; critical care; medical informatics applications; intensive care units

Introduction

Critical care telemedicine, or tele-ICU, is broadly defined as a collaborative, interprofessional care model for critically ill patients where the bedside intensive care unit (ICU) team and patient are networked to a centralized and often remotely located critical care team using telecommunication and computer systems [1,2]. Applications of tele-ICU include quality improvement, continuous monitoring of patients for early warning of deterioration, and varying degrees of clinical decision support, interventions, and consultations [3,4]. Although there exist several tele-ICU models [5,6], we refer to tele-ICU in the context of continuous patient monitoring and subsequent data generation from application of telemedicine in intensive care settings as opposed to more active models involving computer-generated alerts or those with interventions such as audio and video consultations.

Advances in data management infrastructure, biomedical sensors and devices, and computational methods, coupled with the current trend of consolidation of hospitals into large health care delivery systems, provide unique opportunities for not only enhancing tele-ICU capabilities to improve patient, physician, and system-level outcomes but also leveraging tele-ICU data for research and evaluation purposes. The full benefit of the influx of tele-ICU data, however, has yet to be realized.

The objective of this study was to explore opportunities and challenges of using multisite, multimodal data acquired through critical care telemedicine. Using a publicly available tele-ICU database (eICU Collaborative Research Database), we illustrate the quality and potential uses of remote monitoring data [7]. In addition, we present a case study on extraction of multiple respiratory failure patient cohorts to illustrate various strengths and limitations of the database. Specifically, we present 3 patient cohorts—endotracheal intubation patients, patients requiring other noninvasive ventilation therapy, and patients with both invasive and noninvasive treatments in the same visit—and attempt to generate relevant questions for further research.

Methods

The electronic ICU (eICU) database consists of deidentified data collected from patients admitted to adult ICUs between

2014 and 2015. It consists of a wide array of data from admission diagnosis, patient severity scores, standard and custom lab values, nurse charting, physiological data, and treatment records through discharge status. Clinical scores in the database include the Acute Physiology Score (APS) III and the Acute Physiology and Chronic Health Evaluation (APACHE) IV and IVa, both of which are examples of existing instruments that have been widely used in critical care settings for assessment of disease severity and outcome prediction [8].

Hospital data were extracted along with patient demographics, diagnoses, length of stay and mortality outcomes, and treatment records. APACHE IVa severity scores and prediction values were also extracted. Development of respiratory failure cohorts utilized multiple record types in the database that contain respiratory chart and treatment data. The specific cohorts were created using intubation and ventilator-type records. We then attempted to verify patients that required endotracheal intubation or noninvasive respiratory therapy with a redundant record within the database to validate that patients actually required ventilation. For example, one can confidently say that a patient with a record of *endotracheal tube* and a treatment record of *endotracheal tube insertion* was in fact intubated during their ICU stay compared with a patient who has ventilator setting records but no other indication of airway type or noninvasive respiratory therapy.

Data from the eICU database were extracted and preprocessed in Python version 2.7.14 using the Pandas [9] and Seaborn libraries [10], versions 0.23.4 and 0.9.0, respectively. A complete evaluation of all data tables in the eICU database is available [11].

Results

Participant Characteristics

The eICU database consists of 200,859 adult ICU stays at 208 hospitals including 139,367 unique patients with nearly equal numbers of male and female patients. The majority of patients are white. A high-level overview of the database is shown in Figure 1 and additional patient characteristics are presented in Table 1.

Figure 1. Infographic overview of the eICU Collaborative Research Database. ICU: intensive care unit.

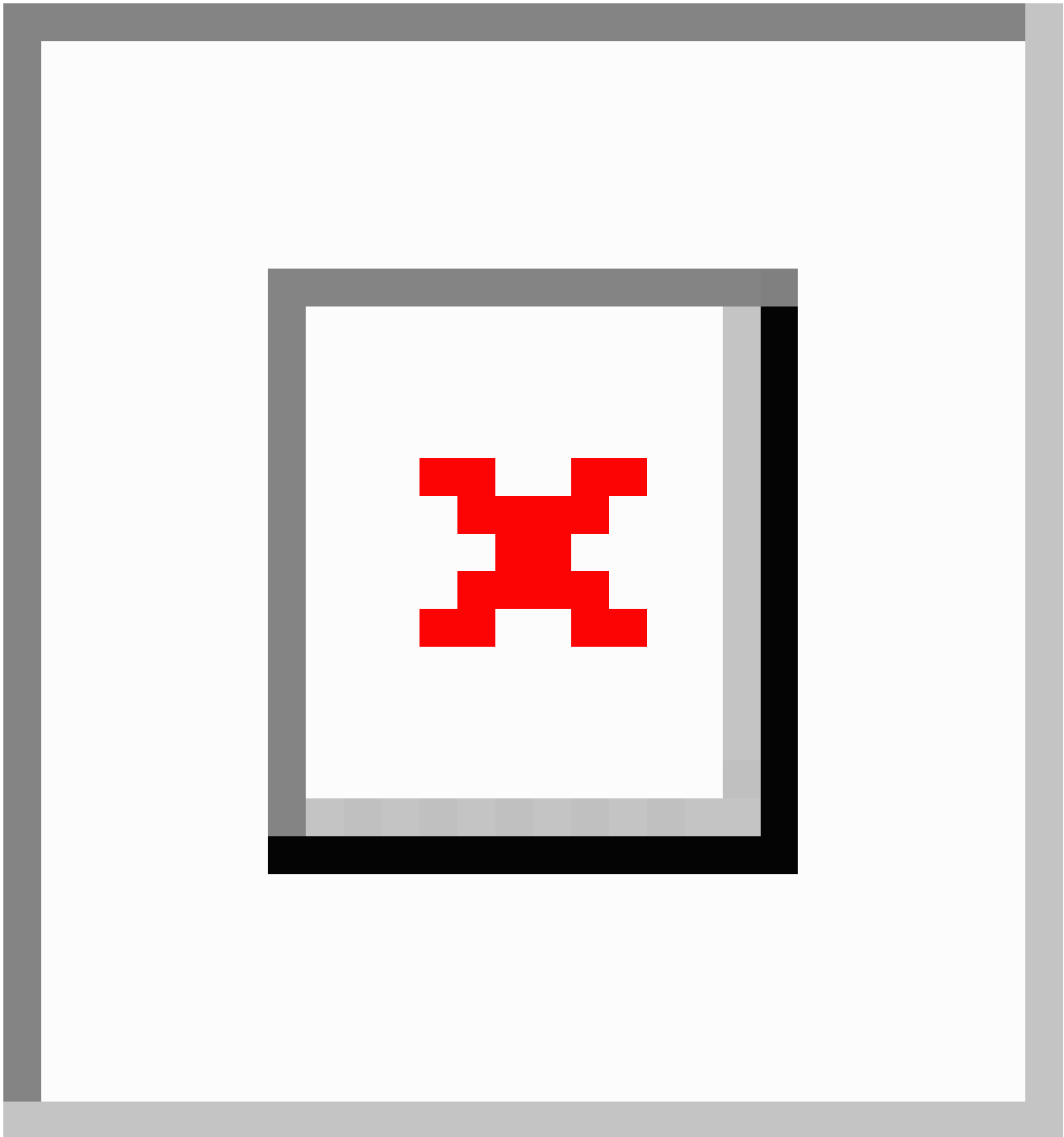


Table 1. Basic patient characteristics in the critical care telemedicine (tele-ICU) database.

Variable	Level	Overall
Unique patients, n	— ^a	139,367
Distinct ICU ^b admissions, n	—	200,859
Age, years, mean (SD)	—	62.1 (16.7)
Gender, n (%)		
Male	—	108,379 (53.96)
Female	—	92,303 (45.95)
Other	—	134 (0.09)
Race, n (%)		
White	—	155,285 (77.31)
African American	—	21,308 (10.61)
Hispanic	—	7464 (3.72)
Asian	—	3270 (1.63)
Native American	—	1700 (0.85)
Unknown or unspecified	—	11,832 (5.90)
ICU length of stay in days, mean (IQR ^c)	ICU	3.00 (2.31)
ICU mortality, % of ICU admissions	ICU	5.79
Hospital length of stay in days, mean (IQR)	Hospital	8.06 (7.04)
Hospital mortality, % of admissions	Hospital	9.24

^aNot applicable.

^bICU: intensive care unit.

^cIQR: interquartile range.

The ICU types covered in the database are wide ranging, with mixed medical-surgical ICU as the most common critical care setting (Figure 2). This is likely because of the configuration and workflow of ICUs within each hospital. The majority of hospitals in the eICU database are primarily nonteaching hospitals across most of the United States (Figure 3).

There were 431 admission diagnoses with several additional diagnosis records in the database that provide context and higher

granularity to the reasons for admission. Patients with cardiovascular conditions accounted for more than 20% of ICU stays, and those with neurological or respiratory illness accounted for nearly 15% of ICU unit stays. Table 2 shows further details on the most frequent admission diagnosis by number of ICU stays and the associated percent of the total visits in the database with corresponding mortality rates and average ICU length of stay.

Figure 2. Frequency of admission to each intensive care unit type within the eICU Collaborative Research Database. ICU: intensive care unit; Med-Surg ICU: medical surgical ICU; CTICU: cardiothoracic ICU; SICU: surgical ICU; CCU-CTICU: coronary care/CTICU ICU; MICU: medical ICU; Neuro ICU: neurological ICU; Cardiac ICU: cardiological ICU; CSICU: cardiac surgery ICU.

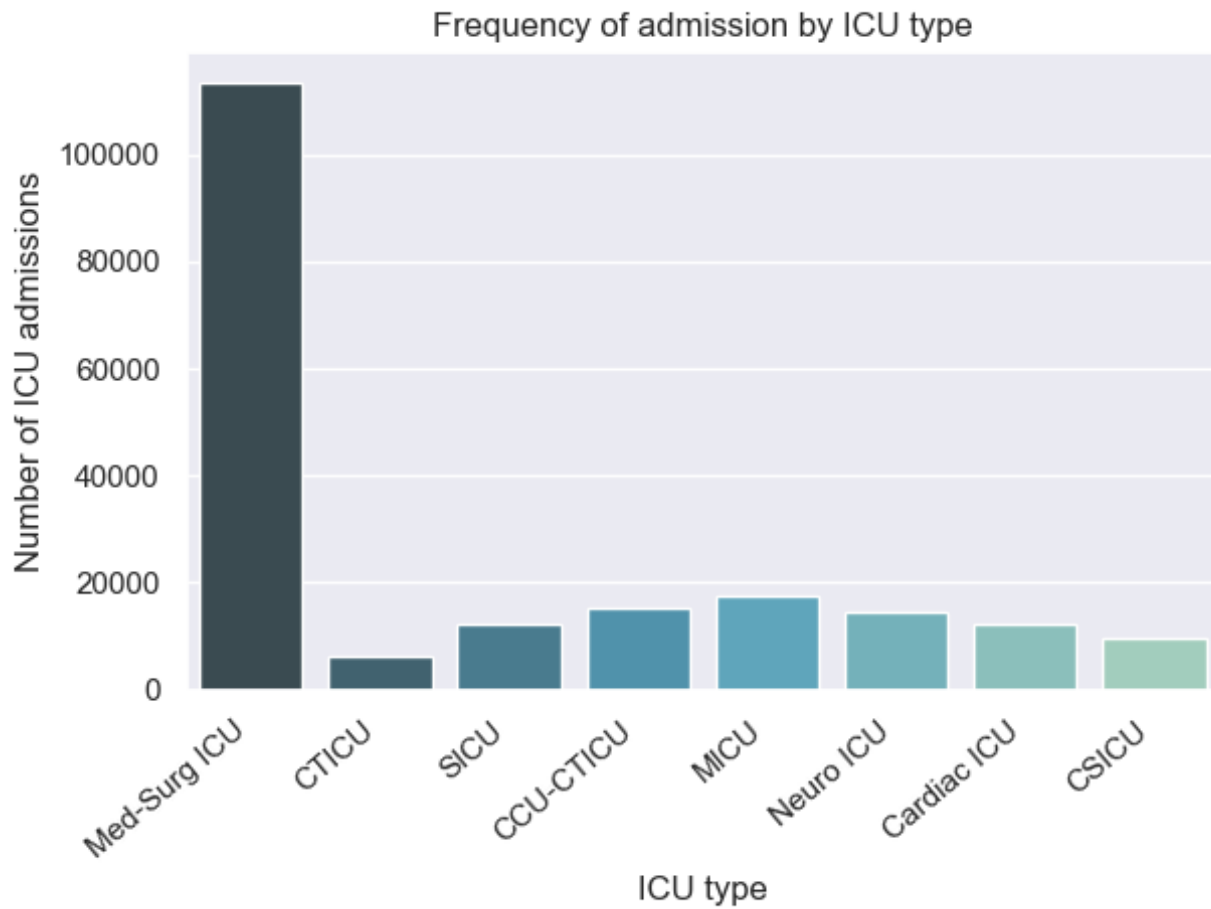


Figure 3. (a) Hospital distribution by size and associated teaching status (b) hospital distribution by United States region and associated teaching status.

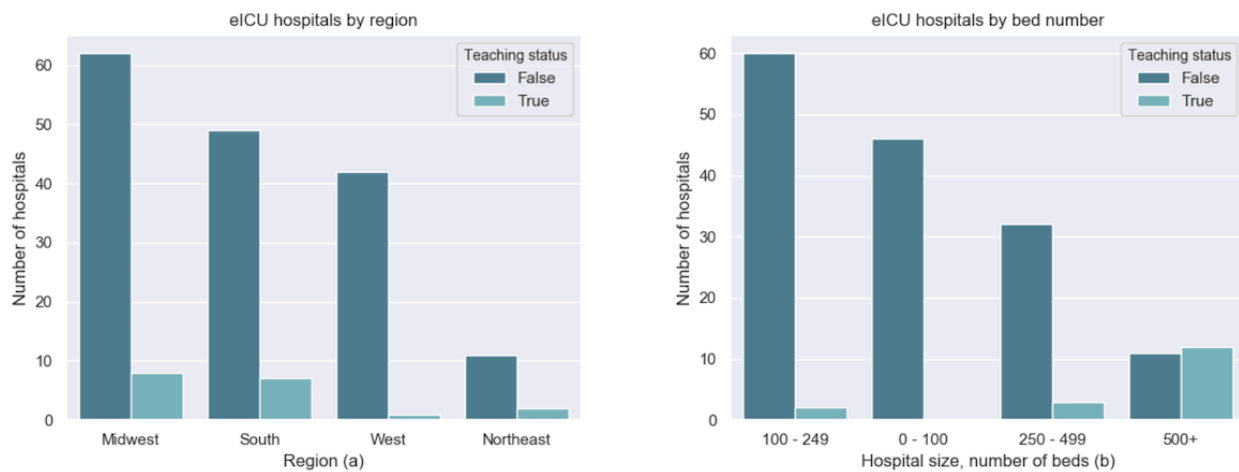


Table 2. Most frequent admission diagnosis categories with corresponding intensive care unit (ICU) mortality rate and average ICU length of stay.

Admission diagnosis name	ICU stays, n (%)	Average length of stay, days	Mortality, n (%)
Cardiovascular	79,560 (20.6)	2.97	4861 (7.33)
Neurologic	31,113 (8.07)	2.83	949 (3.64)
Respiratory	25,813 (6.69)	3.68	1408 (7.00)
Gastrointestinal	17,726 (4.60)	2.95	681 (4.63)
Sepsis, pulmonary	8862 (2.30)	4.31	904 (12.26)
Metabolic or endocrine	8025 (2.08)	1.88	72 (1.06)
Infarction, acute myocardial	7228 (1.87)	2.09	180 (2.93)
Trauma	7136 (1.85)	3.59	303 (5.01)
Cerebrovascular accident or stroke	6647 (1.72)	2.77	290 (5.20)
Congestive heart failure	6617 (1.72)	3.13	302 (5.67)

Table 3. Overview of Acute Physiology Score and Acute Physiology and Chronic Health Evaluation (APACHE) scores in the tele-ICU (critical care telemedicine) database with APACHE IVa predictions.

Variable	Overall	Predicted ^a	Actual
Acute Physiology Score, mean (IQR ^b)	43.63 (27.00)	— ^c	—
APACHE Score, mean (IQR)	55.49 (31.00)	—	—
Intensive care unit (ICU) length of stay in days, mean (IQR)	—	3.87 (3.02)	3.00 (2.31)
ICU mortality, % of ICU admissions ^d	—	5.49	5.79
Hospital length of stay in days, mean (IQR)	—	9.44 (5.88)	8.06 (7.04)
Hospital mortality, % of admissions ^d	—	3.84	9.24

^aPrediction values taken from APACHE version IVa.

^bIQR: interquartile range.

^cN/A: not applicable.

^dPredicted ICU and hospital mortality values are the averages of percent chance of dying of all patients.

Severity and Predictive Scoring Systems

Severity of illness and prognosis are captured in the eICU database as a function of the APACHE IVa score and consists of 288,090 entries. The APACHE evaluation also provides predictions of patient outcomes soon after ICU admission and includes probability of mortality, length of stay, and ventilation days and is used in conjunction with APS. An overview of APS and APACHE scores is presented in Table 3. The distributions of patient severity within the eICU database as a function of APACHE IVa stratified by discharge status of alive or expired is shown in Figure 4.

The APACHE mortality prediction distributions were normalized and segregated by discharge status as shown in Figure 5. This illustrates existing model deficiencies where predicted mortality is not reliable at higher severities [12,13]. Although the predictions for the survivors are reasonably accurate, the predictions for nonsurvivors are not. We include this to illustrate that although predictive models are useful in certain situations, they may not perform well in others because of the dynamics involved or issues with source data [14]. These results are consistent with evaluations of earlier versions of APACHE predictions [15] and are an area of improvement for tele-ICU to provide the best possible decision support for the fast-paced ICU environment.

Figure 4. Kernel density estimate (KDE) of Acute Physiology and Chronic Health Evaluation (APACHE) IVa scores within the eICU Collaborative Research Database stratified by actual intensive care unit mortality outcome.

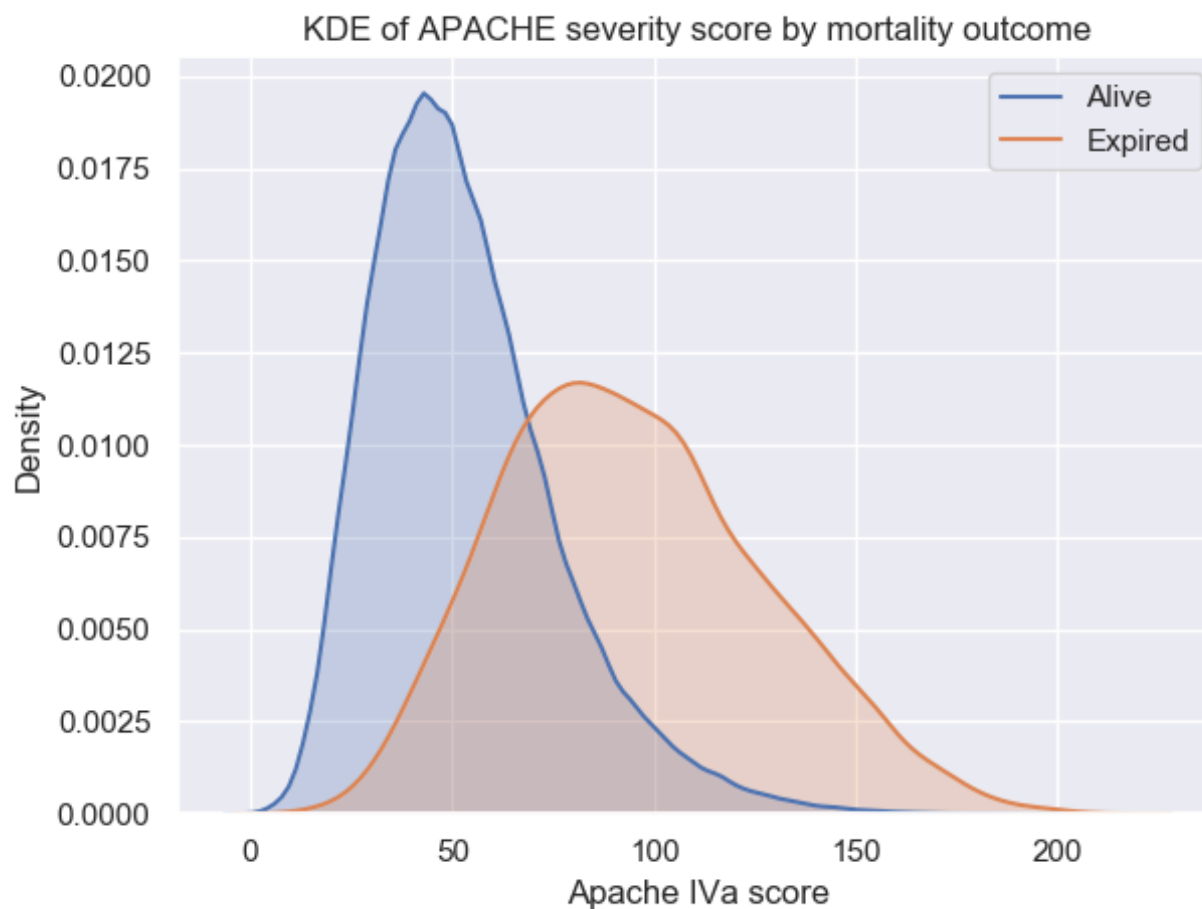
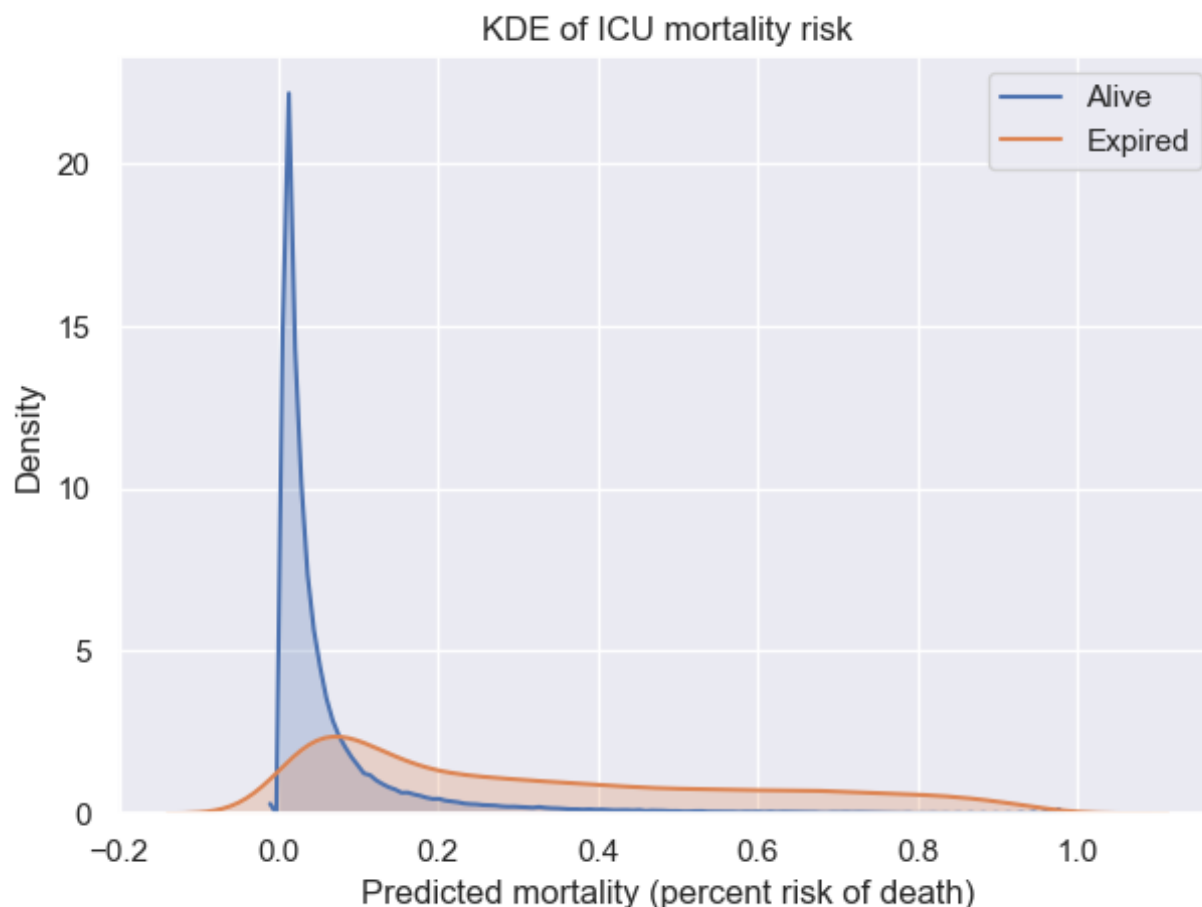


Figure 5. Kernel density estimate (KDE) of predicted hospital mortality in the eICU Collaborative Research Database stratified by actual intensive care unit (ICU) mortality outcome.



Case Study on Respiratory Failure Patients

The selected respiratory failure patient cohorts and corresponding number of patients within each group developed from treatment records are shown in [Figure 6](#). Possible noninvasive ventilation therapy failure was determined using treatment timestamps. Many endotracheal intubation records correspond to continuous positive airway pressure (CPAP) and positive end expiratory pressure (PEEP) records at the same time. However, it is possible that patients with intubation

treatment recorded after CPAP or PEEP treatment required intubation after failure of noninvasive respiratory therapy.

As a demonstration of database coverage and specificity within a particular patient cohort, we selected the 1004 patients that have definitive records of both endotracheal tube insertion and removal. Using the associated treatment time stamps for tube insertion and removal, intubation times were estimated as was the distribution of admission diagnoses across the same cohort ([Figures 7 and 8](#)).

Figure 6. Number of patients with particular respiratory-type treatment records in the eICU database. CPAP: continuous positive airway pressure; PEEP: positive end expiratory pressure.

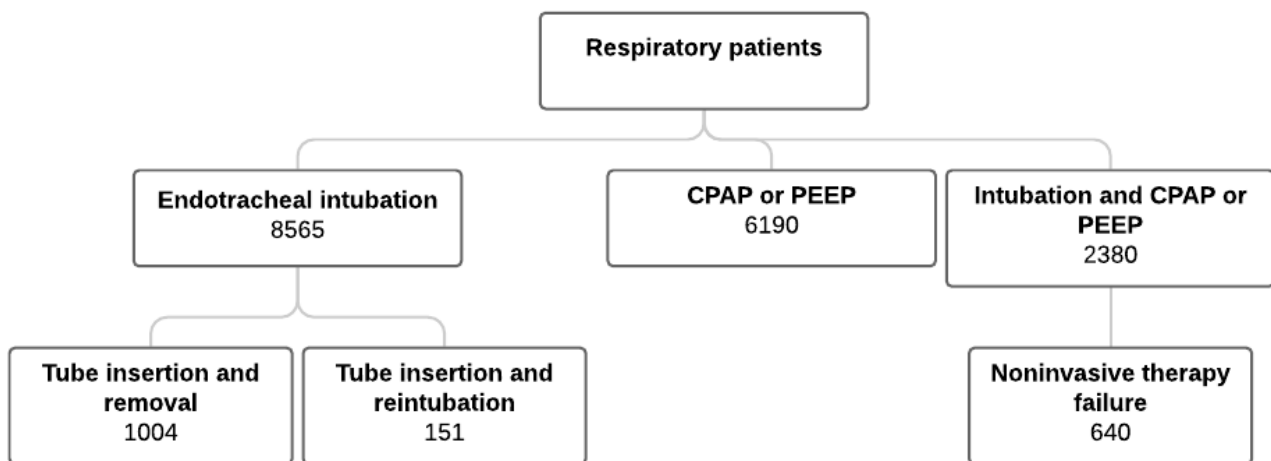


Figure 7. Kernel density estimate (KDE) of intubation times for patients with endotracheal tube insertion and removal.

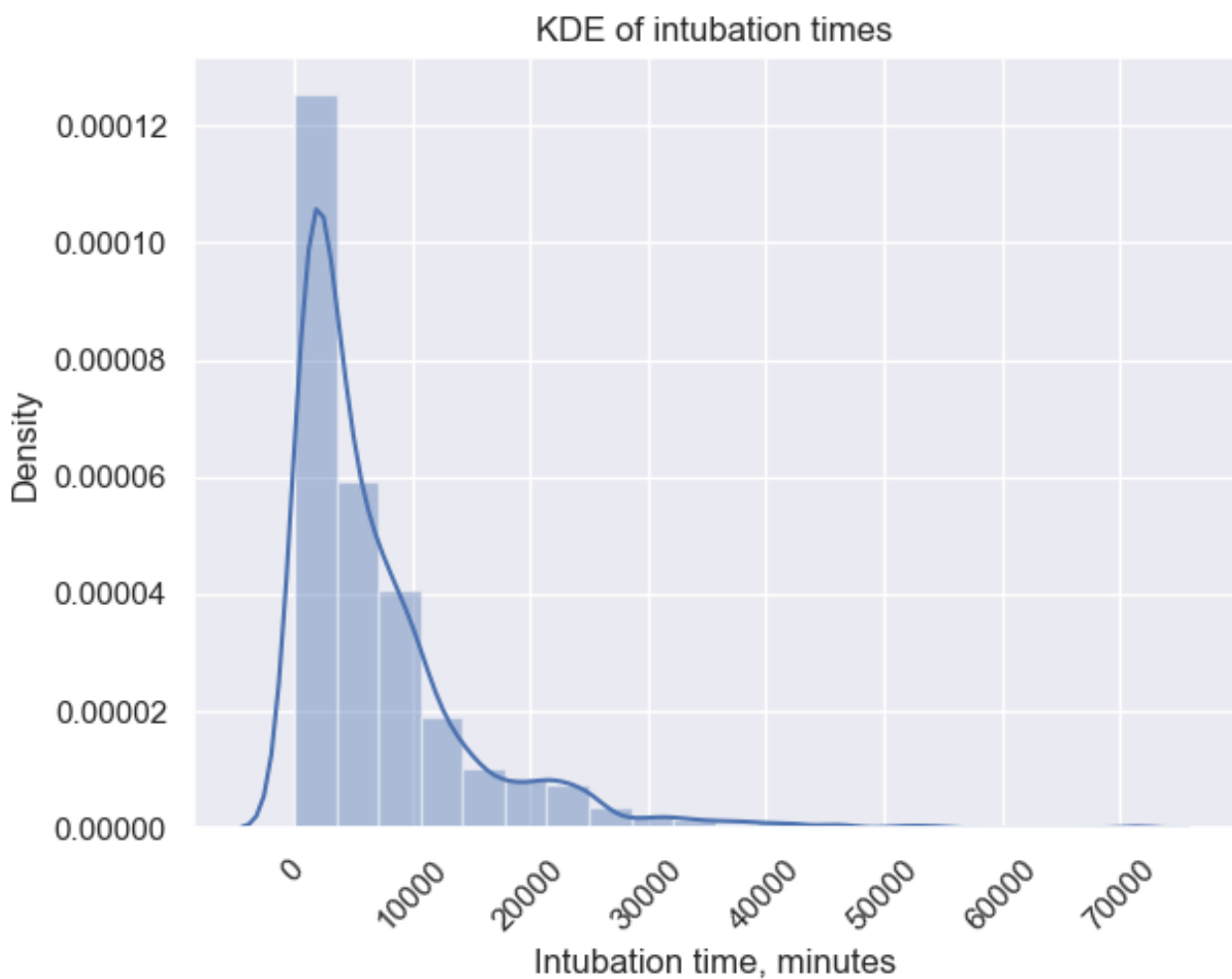
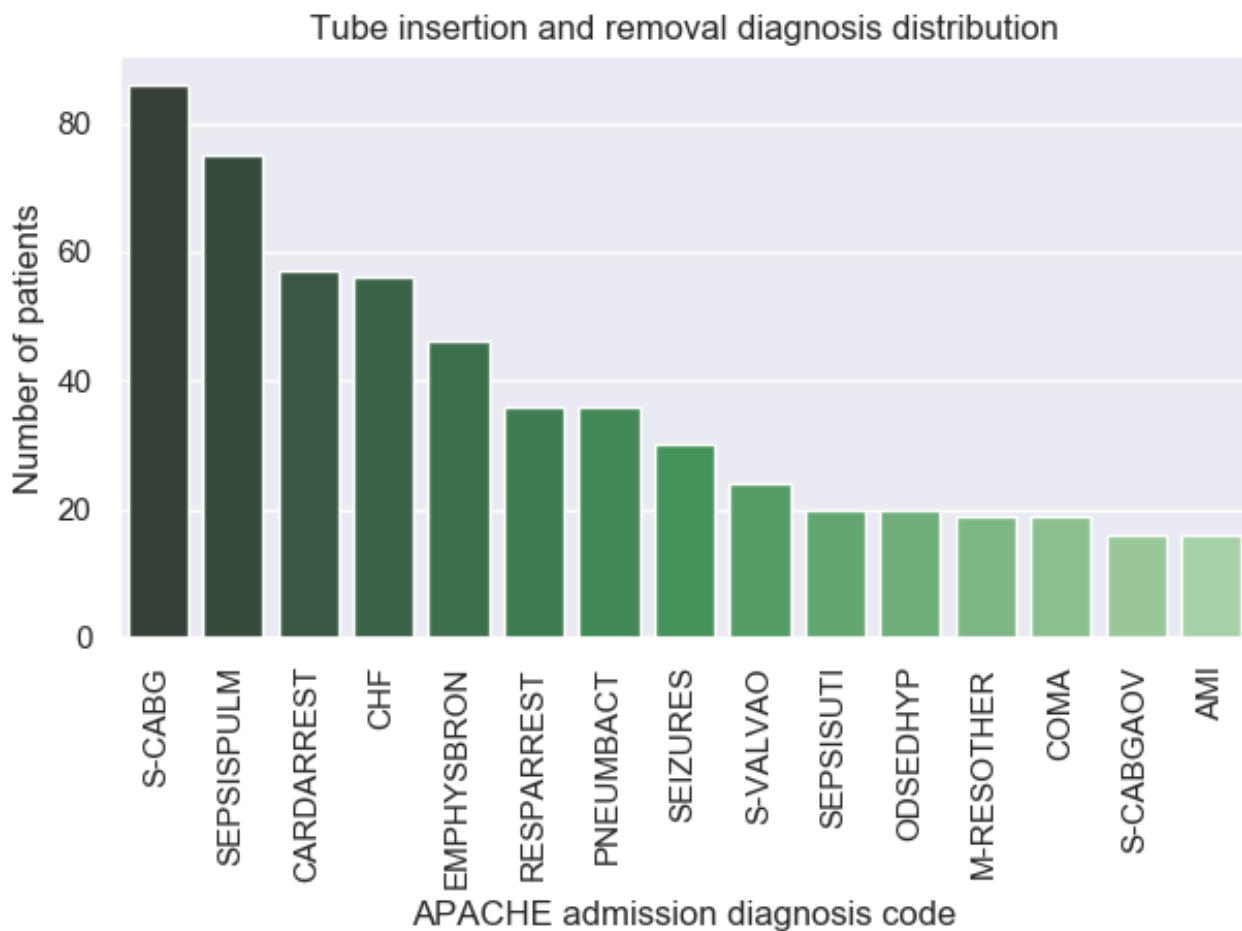


Figure 8. Fifteen most frequent admission diagnoses for patients with record of endotracheal tube insertion and tube removal. AMI: acute myocardial infarction; CABG: coronary artery bypass grafting; CHF: congestive heart failure.



Discussion

Principal Findings

Investigation of the eICU Collaborative Research Database shows a wide range of illnesses from a large number of hospitals that span the continental United States. Organized as a relational database, it is highly versatile for narrowing research focus to specific critical care patient populations, and it allows for robust and generalizable analysis and modeling across multiple institutions and regions. The case study on respiratory failure patients illustrates the potential for cohort discovery and analysis of specific patient subgroups (see Figure 6) using unique identifiers across the database, coupled with the ability to query multiple record types such as treatment records, respiratory, medication, or laboratory data. For example, if using *treatment* records, one would find 8565 unique patients that required endotracheal intubation. If searching for patients with distinct records of both endotracheal tube insertion and removal treatments, the available cohort is limited to 1004 patients. Any combination of these data with other record types may limit or extend cohort size further.

Although the eICU database provides real-world critical care data from a diverse sample of hospitals and practice settings to evaluate interventions, there are some limitations to consider. First, the granularity of the data can be limiting, given the nature

of data collection. For example, despite continuous collection of hemodynamic data, interventions such as tracheal intubation may be recorded with a margin of error because of a requirement for manual entry of events into the electronic medical record. Narrowing the window between when the intubation was performed and when the event was recorded could potentially be accomplished by using drugs associated with intubation. Regardless, this limitation makes studying peri-intubation complications difficult as one does not know whether a hemodynamic decompensation occurred before or after the intubation procedure. In addition, manually entered data could have deviations based on hospital-specific practices and protocol variations.

Second, though the eICU database is considered tele-ICU data, the mode of data collection and the origins of data are not well defined. Specifically, it is not clear which data are generated at the bedside versus the remote unit and by whom. Third, terminology variations across institutions and health information systems pose an additional hurdle. A study of a previous version of eICU data showed discrepancies in standards for laboratory and microbiology data for patients with primary cardiovascular diagnosis [16]. This suggests that cohort discovery on eICU data may also need to be reconfigured based on specific research questions.

Finally, a major caveat to the eICU database is that the absence of a record does not mean an event did not occur. This is true

of other similar databases; however, missing records are exacerbated in the eICU database because of data being from many different hospitals, and not all participating hospitals have interfaces in place to record all data types. Although there are methods for handling data sparsity and missing data [17,18], large quantities of missing data could negate the overall benefit of having a large number of hospitals in the database.

Despite these limitations, the most critical component of future tele-ICU operations and the eICU Collaborative Research Database is that of advanced analytics and clinical decision support. For example, cardiovascular complications arising from traumatic brain injury are common and are linked to increased morbidity and mortality [19]. Generally, monitoring the vital signs of the patient and controlling primary intracranial pathology are effective for proactive prevention of complications. Tele-ICU not only offers continuous display of vitals for remote monitoring but can also serve as a platform to (1) develop, implement, and test clinical and subclinical markers of patient decompensation and other adverse events and (2) further define the role of tele-ICU in improving the precision of electronic alerts [20]. For example, as alarm desensitization creates additional risk, much of the monitoring and resolving of alerts can be shifted to the tele-ICU [21].

Other large, publicly available databases, such as the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) database, have been widely used for secondary research purposes. MIMIC includes inpatient critical care data,

spanning over 10 years from a single institution. An evaluation of the MIMIC database highlights the successful role of the database in risk assessment, medical personnel performance evaluation, and supporting development of clinical decision support systems [22]. The eICU database has strong potential for advancing the role of critical care telemedicine as MIMIC has been for bedside, inpatient critical care.

Conclusions

This work, to our knowledge, is the first of its kind to demonstrate the potential and versatility of a publicly available, large, critical care telemedicine database. The ability to extract and analyze wide-ranging patient subgroups from remote monitoring data for secondary research is one of the key strengths of such a resource. As highlighted through the case study on respiratory patients, there are some limitations such as data provenance and sparsity, which are typical of such resources. Nonetheless, tele-ICU data are particularly useful to catalyze efforts around developing robust clinical decision support systems for critical care that can be distributed between bedside and remote care teams, as well as identifying specific patient populations and associated clinical events that would be appropriate for such *distributed* care. Secondary insults, in particular, stand to benefit from remote monitoring and advanced analytical support because they are highly time sensitive and potentially reversible in critically ill patients, if mitigated promptly. The eICU database and the resulting tools will also inform coordination of care, intensivist coverage, and the overall process of critical care telemedicine.

Acknowledgments

This work was supported in part by the Office of Research, Discovery, Innovation at the University of Arizona, the National Science Foundation under grant #1838745, and the National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 5T32HL007955.

Conflicts of Interest

None declared.

References

1. Goran SF. A second set of eyes: an introduction to tele-ICU. *Crit Care Nurse* 2010 Aug;30(4):46-55 [FREE Full text] [doi: [10.4037/ccn2010283](https://doi.org/10.4037/ccn2010283)] [Medline: [20675821](https://pubmed.ncbi.nlm.nih.gov/20675821/)]
2. Davis TM, Barden C, Dean S, Gavish A, Goliash I, Goran S, et al. American Telemedicine Association guidelines for teleICU operations. *Telemed J E Health* 2016 Dec;22(12):971-980. [doi: [10.1089/tmj.2016.0065](https://doi.org/10.1089/tmj.2016.0065)] [Medline: [27508454](https://pubmed.ncbi.nlm.nih.gov/27508454/)]
3. Kahn JM. The use and misuse of ICU telemedicine. *J Am Med Assoc* 2011 Jun 01;305(21):2227-2228. [doi: [10.1001/jama.2011.716](https://doi.org/10.1001/jama.2011.716)] [Medline: [21576623](https://pubmed.ncbi.nlm.nih.gov/21576623/)]
4. Perednia DA, Allen A. Telemedicine technology and clinical applications. *J Am Med Assoc* 1995 Feb 08;273(6):483-488. [doi: [10.1001/jama.1995.03520300057037](https://doi.org/10.1001/jama.1995.03520300057037)] [Medline: [7837367](https://pubmed.ncbi.nlm.nih.gov/7837367/)]
5. Ramnath VR, Khazeni N. Centralized monitoring and virtual consultant models of tele-ICU care: a side-by-side review. *Telemed J E Health* 2014 Oct;20(10):962-971. [doi: [10.1089/tmj.2014.0024](https://doi.org/10.1089/tmj.2014.0024)] [Medline: [25225795](https://pubmed.ncbi.nlm.nih.gov/25225795/)]
6. Wilcox ME, Adhikari NK. The effect of telemedicine in critically ill patients: systematic review and meta-analysis. *Crit Care* 2012 Jul 18;16(4):R127 [FREE Full text] [doi: [10.1186/cc11429](https://doi.org/10.1186/cc11429)] [Medline: [22809335](https://pubmed.ncbi.nlm.nih.gov/22809335/)]
7. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000 Jun 13;101(23):E215-E220. [doi: [10.1161/01.CIR.101.23.e215](https://doi.org/10.1161/01.CIR.101.23.e215)] [Medline: [10851218](https://pubmed.ncbi.nlm.nih.gov/10851218/)]
8. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. *Crit Care* 2010;14(2):207 [FREE Full text] [doi: [10.1186/cc8204](https://doi.org/10.1186/cc8204)] [Medline: [20392287](https://pubmed.ncbi.nlm.nih.gov/20392287/)]

9. van der Walt S, Millman J. SciPy. 2007 May. Proceedings of the 9th Python in Science Conference URL: <http://conference.scipy.org/proceedings/scipy2010/pdfs/proceedings.pdf> [accessed 2019-01-12] [WebCite Cache ID 75N3pWhAF]
10. Waskom M, Botvinnik O, Hobson P, Cole JB, Halchenko Y, Hoyer S, et al. Zenodo. 2018 Jul 16. seaborn: v0.5.0 (November 2014) URL: <https://zenodo.org/record/12710#.XD5FRFwzbIV> [accessed 2019-01-10] [WebCite Cache ID 75KgXqu5y]
11. Pollard TJ, Johnson AE, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data* 2018 Dec 11;5:180178 [FREE Full text] [doi: [10.1038/sdata.2018.178](https://doi.org/10.1038/sdata.2018.178)] [Medline: [30204154](https://pubmed.ncbi.nlm.nih.gov/30204154/)]
12. Wong LS, Young JD. A comparison of ICU mortality prediction using the APACHE II scoring system and artificial neural networks. *Anaesthesia* 1999 Nov;54(11):1048-1054 [FREE Full text] [doi: [10.1046/j.1365-2044.1999.01104.x](https://doi.org/10.1046/j.1365-2044.1999.01104.x)] [Medline: [10540093](https://pubmed.ncbi.nlm.nih.gov/10540093/)]
13. Einav L, Finkelstein A, Mullainathan S, Obermeyer Z. Predictive modeling of US health care spending in late life. *Science* 2018 Dec 29;360(6396):1462-1465 [FREE Full text] [doi: [10.1126/science.aar5045](https://doi.org/10.1126/science.aar5045)] [Medline: [29954980](https://pubmed.ncbi.nlm.nih.gov/29954980/)]
14. Balkan B, Essay P, Subbian V. Evaluating ICU clinical severity scoring systems and machine learning applications: APACHE IV/IVa case study. In: Proc IEEE Eng Med Biol Soc. 2018 Jul Presented at: IEEE Eng Med Biol Soc; July 18-21, 2018; Honolulu, HI, USA p. 4073-4076. [doi: [10.1109/EMBC.2018.8513324](https://doi.org/10.1109/EMBC.2018.8513324)]
15. Goldhill DR, Sumner A. APACHE II, data accuracy and outcome prediction. *Anaesthesia* 1998 Oct;53(10):937-943 [FREE Full text] [doi: [10.1046/j.1365-2044.1998.00534.x](https://doi.org/10.1046/j.1365-2044.1998.00534.x)] [Medline: [9893535](https://pubmed.ncbi.nlm.nih.gov/9893535/)]
16. Chronaki C, Shahin A, Mark R. Designing reliable cohorts of cardiac patients across MIMIC and eICU. *Comput Cardiol* 2015;42:189-192 [FREE Full text] [doi: [10.1109/CIC.2015.7408618](https://doi.org/10.1109/CIC.2015.7408618)] [Medline: [27774488](https://pubmed.ncbi.nlm.nih.gov/27774488/)]
17. Waljee AK, Mukherjee A, Singal AG, Zhang Y, Warren J, Balis U, et al. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* 2013 Aug 01;3(8) [FREE Full text] [doi: [10.1136/bmjopen-2013-002847](https://doi.org/10.1136/bmjopen-2013-002847)] [Medline: [23906948](https://pubmed.ncbi.nlm.nih.gov/23906948/)]
18. Vesin A, Azoulay E, Ruckly S, Vignoud L, Rusinová K, Benoit D, et al. Reporting and handling missing values in clinical studies in intensive care units. *Intensive Care Med* 2013 Aug;39(8):1396-1404. [doi: [10.1007/s00134-013-2949-1](https://doi.org/10.1007/s00134-013-2949-1)] [Medline: [23685609](https://pubmed.ncbi.nlm.nih.gov/23685609/)]
19. Manley G, Knudson MM, Morabito D, Damron S, Erickson V, Pitts L. Hypotension, hypoxia, and head injury: frequency, duration, and consequences. *Arch Surg* 2001 Oct;136(10):1118-1123. [doi: [10.1001/archsurg.136.10.1118](https://doi.org/10.1001/archsurg.136.10.1118)] [Medline: [11585502](https://pubmed.ncbi.nlm.nih.gov/11585502/)]
20. Celi LA, Hassan E, Marquardt C, Breslow M, Rosenfeld B. The eICU: it's not just telemedicine. *Crit Care Med* 2001 Aug;29(8 Suppl):N183-N189. [doi: [10.1097/00003246-200108001-00007](https://doi.org/10.1097/00003246-200108001-00007)] [Medline: [11496041](https://pubmed.ncbi.nlm.nih.gov/11496041/)]
21. Johnson KR, Hagadorn JI, Sink DW. Alarm safety and alarm fatigue. *Clin Perinatol* 2017 Sep;44(3):713-728. [doi: [10.1016/j.clp.2017.05.005](https://doi.org/10.1016/j.clp.2017.05.005)] [Medline: [28802348](https://pubmed.ncbi.nlm.nih.gov/28802348/)]
22. Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LW, Moody G, et al. Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. *Crit Care Med* 2011 May;39(5):952-960 [FREE Full text] [doi: [10.1097/CCM.0b013e31820a92c6](https://doi.org/10.1097/CCM.0b013e31820a92c6)] [Medline: [21283005](https://pubmed.ncbi.nlm.nih.gov/21283005/)]

Abbreviations

- APACHE:** Acute Physiology and Chronic Health Evaluation
APS: Acute Physiology Score
CPAP: continuous positive airway pressure
eICU: electronic intensive care unit
ICU: intensive care unit
IQR: interquartile range
MIMIC: Multiparameter Intelligent Monitoring in Intensive Care
PEEP: positive end expiratory pressure
tele-ICU: critical care telemedicine

Edited by G Eysenbach; submitted 02.12.18; peer-reviewed by T Abdulai, T Aslanidis; comments to author 26.12.18; revised version received 29.12.18; accepted 29.12.18; published 24.01.19.

Please cite as:

Essay P, Shahin TB, Balkan B, Mosier J, Subbian V

The Connected Intensive Care Unit Patient: Exploratory Analyses and Cohort Discovery From a Critical Care Telemedicine Database
JMIR Med Inform 2019;7(1):e13006

URL: <http://medinform.jmir.org/2019/1/e13006/>

doi: [10.2196/13006](https://doi.org/10.2196/13006)

PMID: [30679148](https://pubmed.ncbi.nlm.nih.gov/30679148/)

©Patrick Essay, Tala B Shahin, Baran Balkan, Jarrod Mosier, Vignesh Subbian. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 24.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Predicting Appropriate Hospital Admission of Emergency Department Patients with Bronchiolitis: Secondary Analysis

Gang Luo¹, PhD; Bryan L Stone², MSc, MD; Flory L Nkoy², MSc, MPH, MD; Shan He³, PhD; Michael D Johnson², MD

¹Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

²Department of Pediatrics, University of Utah, Salt Lake City, UT, United States

³Care Transformation, Intermountain Healthcare, Salt Lake City, UT, United States

Corresponding Author:

Gang Luo, PhD

Department of Biomedical Informatics and Medical Education

University of Washington

Building C, Box 358047

850 Republican Street

Seattle, WA, 98195

United States

Phone: 1 206 221 4596

Fax: 1 206 221 2671

Email: gangluo@cs.wisc.edu

Abstract

Background: In children below the age of 2 years, bronchiolitis is the most common reason for hospitalization. Each year in the United States, bronchiolitis causes 287,000 emergency department visits, 32%-40% of which result in hospitalization. Due to a lack of evidence and objective criteria for managing bronchiolitis, clinicians often make emergency department disposition decisions on hospitalization or discharge to home subjectively, leading to large practice variation. Our recent study provided the first operational definition of appropriate hospital admission for emergency department patients with bronchiolitis and showed that 6.08% of emergency department disposition decisions for bronchiolitis were inappropriate. An accurate model for predicting appropriate hospital admission can guide emergency department disposition decisions for bronchiolitis and improve outcomes, but has not been developed thus far.

Objective: The objective of this study was to develop a reasonably accurate model for predicting appropriate hospital admission.

Methods: Using Intermountain Healthcare data from 2011-2014, we developed the first machine learning classification model to predict appropriate hospital admission for emergency department patients with bronchiolitis.

Results: Our model achieved an accuracy of 90.66% (3242/3576, 95% CI: 89.68-91.64), a sensitivity of 92.09% (1083/1176, 95% CI: 90.33-93.56), a specificity of 89.96% (2159/2400, 95% CI: 88.69-91.17), and an area under the receiver operating characteristic curve of 0.960 (95% CI: 0.954-0.966). We identified possible improvements to the model to guide future research on this topic.

Conclusions: Our model has good accuracy for predicting appropriate hospital admission for emergency department patients with bronchiolitis. With further improvement, our model could serve as a foundation for building decision-support tools to guide disposition decisions for children with bronchiolitis presenting to emergency departments.

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.5155

(*JMIR Med Inform* 2019;7(1):e12591) doi:[10.2196/12591](https://doi.org/10.2196/12591)

KEYWORDS

bronchiolitis; appropriate hospital admission; emergency department; predictive model; machine learning

Introduction

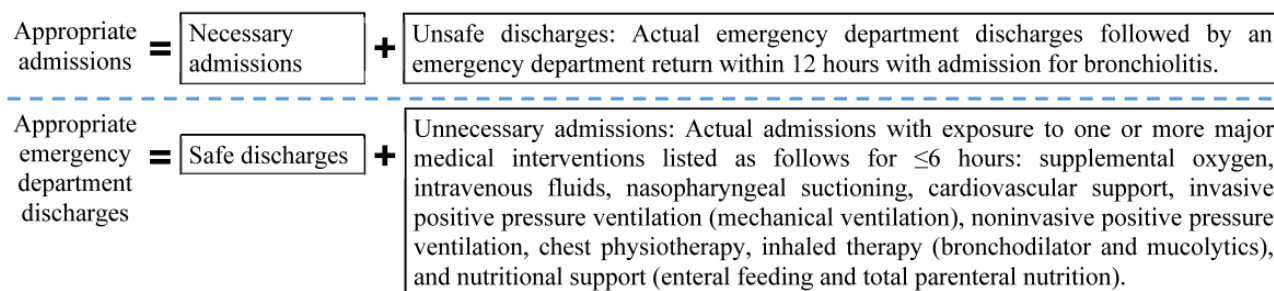
Bronchiolitis refers to inflammation of the bronchioles, the smallest air passages in the lungs, mainly seen in children below the age of 2 years [1]. More than one-third of children in the United States have been diagnosed with bronchiolitis by the age of 2 years [1]. In children below the age of 2 years, bronchiolitis is responsible for 16% of the hospitalizations and is the most common reason for hospitalization [2-5]. In the United States, bronchiolitis annually leads to approximately 287,000 emergency department (ED) visits [6], 128,000 hospitalizations [2], and US \$1.73 billion in total inpatient costs (2009) [2].

About 32%-40% of ED visits for bronchiolitis result in hospitalization [7-9]. Current clinical guidelines for bronchiolitis [10,11] acknowledge that due to a lack of evidence and objective criteria for managing bronchiolitis, clinicians often make ED disposition decisions of hospitalization or discharge to home subjectively [4,12]. This uncertainty in bronchiolitis management leads to large practice variation [3,12-23], increased iatrogenic risk, suboptimal outcomes, and wasted healthcare resources resulting from unnecessary admissions and unsafe discharges [15,21,24]. Approximately 10% of infants with bronchiolitis experience adverse events during hospital stay [25]. By examining the distributions of multiple relevant attributes of ED visits for bronchiolitis and using a data-driven method to determine two threshold values, we recently developed the first operational definition of appropriate hospital admission for ED patients with bronchiolitis [26]. Appropriate

admissions cover both necessary admissions (actual admissions that are necessary) and unsafe discharges (Figure 1). Appropriate ED discharges cover both safe discharges and unnecessary admissions. Unsafe discharges are defined based on early ED returns. Unnecessary admissions are defined based on brief exposure to certain major medical interventions (Figure 1). Brief exposure was defined as exposure of ≤ 6 hours, with the threshold value of 6 hours chosen conservatively based on the median duration of major medical interventions received by a subset of patients who tended to have been admitted unnecessarily. Based on the operational definition, we showed that 6.08% of ED disposition decisions for bronchiolitis were inappropriate [26].

Thus far, several models have been built for predicting hospital admission in ED patients with bronchiolitis [7-9,27-29]. As our review paper [30] pointed out, these models have low accuracy and incorrectly assume that actual ED disposition decisions are always appropriate. An accurate model for predicting appropriate hospital admission can guide ED disposition decisions for bronchiolitis and improve outcomes. This model, which is yet to be built, would be particularly useful for less experienced clinicians, including junior clinicians and those in general practice who attend to children infrequently [31]. The objective of this study was to build the first model to predict appropriate hospital admission for ED patients with bronchiolitis. The dependent variable of the appropriate ED disposition decision is categorical and has two possible values: appropriate admission and appropriate ED discharge. Accordingly, the model uses clinical and administrative data to conduct binary classification.

Figure 1. The operational definition of appropriate hospital admission for emergency department patients with bronchiolitis.



Methods

Study Design and Ethical Approval

In this study, we performed secondary analysis of retrospective data. The Institutional Review Boards of the University of Washington Medicine, University of Utah, and Intermountain Healthcare reviewed and approved this study and waived the need for informed consent for all patients.

Patient Population

Our patient cohort consisted of children below the age of 2 years who visited the ED for bronchiolitis in 2013-2014 at any of the 22 Intermountain Healthcare hospitals. Intermountain Healthcare is the largest healthcare system in Utah, with 22 hospitals and 185 clinics delivering ~85% of pediatric care in Utah [32]. Similar to our previous paper [26], we adopted the approach

used in Flaherman et al [33-35] to identify as many ED visits for bronchiolitis as possible. This approach included patients with an ED or hospital International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) primary discharge diagnosis code of bronchiolitis or bronchitis (466.x), viral pneumonia (480.x), adenoviral infection (079.0), rhinovirus infection (079.3), respiratory infection due to influenza (487.0 or 487.1), respiratory syncytial virus (079.6), H1N1 influenza (488.1, 488.11, or 488.12), influenza due to identified avian influenza virus (488, 488.0, 488.01, or 488.02), or influenza due to novel influenza A (488.81 or 488.82). Any of these discharge diagnosis codes, rather than only the discharge diagnosis code of bronchiolitis, could be assigned to an ED visit for bronchiolitis. In addition, this approach included all patients with any of the abovementioned codes as a nonprimary diagnosis code, as long as the ICD-9-CM primary diagnosis code was any of the following: apnea (786.03),

shortness of breath (786.05), tachypnea (786.06), wheezing (786.07), other respiratory abnormalities (786.09), cough (786.2), fever (780.60 or 780.61), acute nasopharyngitis (460), acute upper respiratory infections (465.x), other specified viral infection (079.89), urinary tract infection (599.0), pneumonia unspecified organism (486), unspecified viral infection (079.99), volume depletion (276.5x), or respiratory failure (518.81 or 518.82) [26]. The ED visits for bronchiolitis captured by this approach in 2013-2014 are the focus of our study.

Data Set

From Intermountain Healthcare's enterprise data warehouse, we extracted a clinical and administrative data set containing information of our patient cohort's inpatient stays, ED visits, and outpatient visits at Intermountain Healthcare in 2011-2014. Our patient cohort included children below the age of 2 years who visited the Intermountain Healthcare ED for bronchiolitis in 2013-2014. By starting the data set in 2011, we ensured that for each ED visit by a target patient in 2013-2014, the data set included the patient's complete prior medical history recorded within Intermountain Healthcare and necessary for computing features (also known as independent variables).

Features

The 35 candidate patient features fall into two disjoint categories. *Category 1* includes all known predictors of hospital admission in ED patients with bronchiolitis, which were consistently recorded at Intermountain Healthcare facilities and available as structured attributes in our data set [30,31]. These 15 predictors are age in days, gender, heart rate, respiratory rate, peripheral capillary oxygen saturation (SpO₂), temperature, coinfection, rhinovirus infection, enterovirus infection, history of bronchopulmonary dysplasia, history of eczema, prior intubation, prior hospitalization, prematurity, and dehydration. For any vital sign that was recorded more than once during the ED visit, we used its last value as its feature value. Among all recorded values, the last value most closely reflected the patient's status at the time of ED disposition.

Category 2 consists of 20 features suggested by our team's clinical experts BLS, MDJ, and FLN: race, ethnicity, insurance category (public, private, or self-paid or charity), the ED visit's acuity level (resuscitation, emergent, urgent, semiurgent, or nonurgent), chief complaint, number of consults during the ED visit, number of laboratory tests ordered during the ED visit, number of radiology studies ordered during the ED visit, number of X-rays ordered during the ED visit, length of ED stay in minutes, hour of ED disposition, whether the patient is up-to-date with his/her immunizations, diastolic blood pressure, systolic blood pressure, weight, wheezing (none, expiratory, inspiratory and expiratory, or diminished breath sounds), retractions (none, one location, two locations, or three or more locations), respiratory syncytial virus infection, language barrier to learning, and whether the patient has any other barrier to learning. For either attribute of wheezing and retractions that was recorded more than once during the ED visit, we used its

last value as its feature value. Among all recorded values, the last value most closely reflected the patient's status at ED disposition time.

Based on the timestamp, all candidate features were available as structured attributes in our data set before the time of ED disposition. We used these features to build predictive models.

Data Analysis

Data Preparation

For each ED visit by a patient below the age of 2 years for bronchiolitis in 2013-2014, we used our previously developed operational definition of appropriate admission [26] (Figure 1) to compute the dependent variable's value. For each numerical feature, we examined the data distribution, used the upper and lower bounds given by our team's ED expert MDJ to identify invalid values, and replaced each invalid value with a null value. All temperatures <80°F or >110°F, all weights >50 pounds, all systolic blood pressure values of 0, all SpO₂ values >100%, all respiratory rates >120 breaths/minute, and all heart rates <30 or >300 beats/minute were regarded as physiologically impossible and invalid. To ensure that all data were on the same scale, we standardized each numerical feature by first subtracting its mean and then dividing by its SD. We focused on 2 years of data for ED visits for bronchiolitis (2013-2014). Data from the first year (2013) were used to train predictive models. Data from the second year (2014) were used to evaluate model performance, reflecting use in practice.

Performance Metrics

As shown in Table 1 and the formulas below, we used six standard metrics to measure model performance: accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic curve (AUC). For instance, false negative (FN) is the number of appropriate admissions that the model incorrectly predicts to be ED discharges. Sensitivity measures the proportion of appropriate admissions that the model identifies. Specificity measures the proportion of appropriate ED discharges that the model identifies.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{PPV} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{NPV} = \text{TN} / (\text{TN} + \text{FN})$$

TP is true positive, TN is true negative, and FP is false negative.

For the six performance metrics, we conducted 1000-fold bootstrap analysis [36] to compute their 95% CIs. On each bootstrap sample of the 2014 data, we computed our model's performance metrics. For each of the six performance metrics, the 2.5th and 97.5th percentiles in the 1000 bootstrap samples specified its 95% CI.

Table 1. The error matrix.

Class	Appropriate admission	Appropriate emergency department discharge
Predicted admission	True positive	False positive
Predicted emergency department discharge	False negative	True negative

To show the sensitivity-specificity tradeoff, we plotted the receiver operating characteristic curve. The calibration of a model refers to how well the predicted probabilities of appropriate admission match with the fractions of appropriate admissions in subgroups of ED visits for bronchiolitis. To show model calibration, we drew a calibration plot [36]. A perfect calibration curve would coincide with the diagonal line. In addition, we used the Hosmer-Lemeshow goodness-of-fit test [36] to evaluate model calibration.

Classification Algorithms

We used Weka [37], a widely used open-source machine learning and data mining toolkit, to build machine learning classification models. Machine learning studies computer algorithms that learn from data, such as random forest, support vector machine, and neural network, and has won most data science competitions [38]. Weka integrates many commonly used machine learning algorithms and feature-selection techniques. We considered all 39 machine learning classification algorithms in the standard Weka package and adopted our previously developed automatic machine learning model selection method [39] and the training data of 2013 to automatically select the algorithm, feature-selection technique, and hyperparameter values among all the applicable ones. In a machine learning algorithm, hyperparameters are the parameters whose values are traditionally manually set by the machine learning software user before model training. An example of a hyperparameter is the number of decision trees used in a random forest classifier. Our automatic machine learning model selection method [39] uses the Bayesian optimization (also known as response surface) methodology to automatically explore numerous combinations of algorithm, feature-selection

technique, and hyperparameter values and performs three-fold cross-validation to select the final combination maximizing the AUC. Compared to the other five performance metrics—accuracy, sensitivity, specificity, PPV, and NPV—AUC has the advantage of not relying on the cutoff threshold for deciding between predicted admission and predicted ED discharge.

Demographic and Clinical Characteristics of the Patient Cohort

Tables 2 and 3 show the demographic and clinical characteristics of our patient cohort: children below the age of 2 years who visited the ED for bronchiolitis in 2013 and 2014, respectively. The characteristics are mostly similar between both years. About 40.78% (1640/4022) and 38.26% (1368/3576) of ED visits for bronchiolitis ended in hospitalization in 2013 and 2014, respectively. About 35.80% (1440/4022) and 32.89% (1176/3576) of ED visits for bronchiolitis were deemed to be appropriate hospital admissions in 2013 and 2014, respectively.

Based on the χ^2 two-sample test, for the 2013 data, the ED visits discharged to home and those ending in hospitalization showed the same distribution for gender ($P=.49$) and different distributions for race ($P<.001$), ethnicity ($P=.01$), and insurance category ($P<.001$). For the 2014 data, the ED visits discharged to home and those ending in hospitalization showed the same distribution for gender ($P=.94$) and race ($P=.61$) and different distributions for ethnicity ($P<.001$) and insurance category ($P<.001$). Based on the Cochran-Armitage trend test [41], for both the 2013 and 2014 data, the ED visits discharged to home and those ending in hospitalization showed different distributions for age ($P<.001$).

Table 2. Demographic and clinical characteristics of children under the age of 2 years who visited the emergency department at Intermountain Healthcare hospitals for bronchiolitis in 2013.

Characteristic	Emergency department visits (N=4022), n (%)	Emergency department visits discharged to home (N=2382), n (%)	Emergency department visits ending in hospitalization (N=1640), n (%)
Age			
<2 months	518 (12.88)	211 (8.86)	307 (18.72)
2 to <12 months	2424 (60.27)	1498 (62.89)	926 (56.46)
12 to 24 months	1080 (26.85)	673 (28.25)	407 (24.82)
Gender			
Male	2369 (58.90)	1414 (59.36)	955 (58.23)
Female	1653 (41.10)	968 (40.64)	685 (41.77)
Race			
American Indian or Alaska native	51 (1.27)	26 (1.09)	25 (1.52)
Asian	49 (1.22)	20 (0.84)	29 (1.77)
Black or African American	124 (3.08)	78 (3.27)	46 (2.80)
Native Hawaiian or other Pacific Islander	321 (7.98)	160 (6.72)	161 (9.82)
White	2940 (73.10)	1784 (74.90)	1156 (70.49)
Unknown or not reported	537 (13.35)	314 (13.18)	223 (13.60)
Ethnicity			
Hispanic	1321 (32.84)	826 (34.68)	495 (30.18)
Non-Hispanic	2687 (66.81)	1549 (65.03)	1138 (69.39)
Unknown or not reported	14 (0.35)	7 (0.29)	7 (0.43)
Insurance			
Private	2436 (60.57)	1338 (56.17)	1098 (66.95)
Public	1422 (35.36)	933 (39.17)	489 (29.82)
Self-paid or charity	164 (4.08)	111 (4.66)	53 (3.23)
Asthma	207 (5.15)	72 (3.02)	135 (8.23)
Chronic complex condition [40]	296 (7.36)	60 (2.52)	236 (14.39)

Table 3. Demographic and clinical characteristics of children under the age of 2 years who visited the emergency department at Intermountain Healthcare hospitals for bronchiolitis in 2014.

Characteristic	Emergency department visits (N=3576), n (%)	Emergency department visits discharged to home (N=2208), n (%)	Emergency department visits ending in hospitalization (N=1368), n (%)
Age			
<2 months	454 (12.70)	186 (8.42)	268 (19.59)
2 to <12 months	2079 (58.14)	1379 (62.45)	700 (51.17)
12 to 24 months	1043 (29.17)	643 (29.12)	400 (29.24)
Gender			
Male	2059 (57.58)	1273 (57.65)	786 (57.46)
Female	1517 (42.42)	935 (42.35)	582 (42.54)
Race			
American Indian or Alaska Native	47 (1.31)	31 (1.40)	16 (1.17)
Asian	68 (1.90)	40 (1.81)	28 (2.05)
Black or African American	104 (2.91)	70 (3.17)	34 (2.49)
Native Hawaiian or other Pacific Islander	284 (7.94)	180 (8.15)	104 (7.60)
White	2795 (78.16)	1708 (77.36)	1087 (79.46)
Unknown or not reported	278 (7.77)	179 (8.11)	99 (7.24)
Ethnicity			
Hispanic	1071 (29.95)	727 (32.93)	344 (25.15)
Non-Hispanic	2484 (69.46)	1464 (66.30)	1020 (74.56)
Unknown or not reported	21 (0.59)	17 (0.77)	4 (0.29)
Insurance			
Private	2175 (60.82)	1241 (56.20)	934 (68.27)
Public	1256 (35.12)	860 (38.95)	396 (28.95)
Self-paid or charity	145 (4.05)	107 (4.85)	38 (2.78)
Asthma	210 (5.87)	67 (3.03)	143 (10.45)
Chronic complex condition [40]	252 (7.05)	43 (1.94)	209 (15.28)

Results

Our automatic machine learning model selection method [39] chose the random forest classification algorithm. Random forest can naturally handle missing feature values. Our model was built using this algorithm and the 33 features shown in Table 4. These features are sorted in descending order of their importance values, which were automatically computed by the random forest algorithm in Weka based on average impurity decrease. In general, the features related to the patient's history are ranked lower than those reflecting the patient's status in the current ED visit. This intuitively makes medical sense. Two candidate patient features—ethnicity and the ED visit's acuity level—were not used in our model because they did not increase the model's accuracy.

Figure 2 shows the receiver operating characteristic curve of our model. Weka uses 50% as its default probability cutoff threshold for making binary classifications. Table 5 shows the error matrix of our model. Table 6 compares our model and the

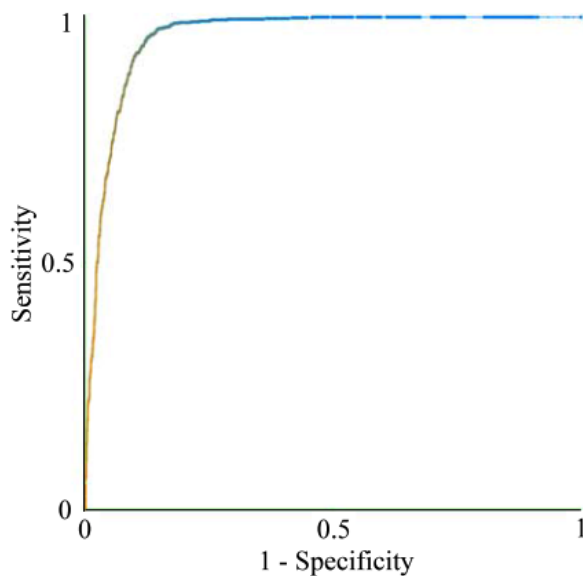
ED clinician's disposition decision. Our model achieved an accuracy of 90.66% (3242/3576; 95% CI: 89.68-91.64), a sensitivity of 92.09% (1083/1176; 95% CI: 90.33-93.56), a specificity of 89.96% (2159/2400; 95% CI: 88.69-91.17), an AUC of 0.960 (95% CI: 0.954-0.966), a PPV of 81.80% (1083/1324; 95% CI: 79.67-83.80), and an NPV of 95.87% (2159/2252; 95% CI: 95.00-96.65). If we removed the insurance category feature, our model achieved a lower accuracy of 90.32% (3230/3576; 95% CI: 89.37-91.28), a lower sensitivity of 90.22% (1061/1176; 95% CI: 88.30-91.79), a specificity of 90.38% (2169/2400; 95% CI: 89.15-91.57), an AUC of 0.960 (95% CI: 0.955-0.966), a PPV of 82.12% (1061/1292; 95% CI: 79.94-84.15), and a lower NPV of 94.97% (2169/2284; 95% CI: 93.97-95.78). In comparison, the ED clinician's disposition decision achieved an accuracy of 93.68% (3350/3576; 95% CI: 92.87-94.49), a sensitivity of 98.55% (1159/1176; 95% CI: 97.85-99.24), a specificity of 91.29% (2191/2400; 95% CI: 90.05-92.46), an AUC of 0.949 (95% CI: 0.942-0.956), a PPV of 84.72% (1159/1368; 95% CI: 82.83-86.69), and an NPV of 99.23% (2191/2208; 95% CI: 98.86-99.59).

Table 4. Features used in our model and their importance.

Feature	Importance based on average impurity decrease
Hour of ED ^a disposition	0.42
Age in days	0.40
Whether the patient has any other barrier to learning	0.39
Length of ED stay in minutes	0.38
Number of laboratory tests ordered during the ED visit	0.37
Heart rate	0.37
Diastolic blood pressure	0.36
Gender	0.35
Temperature	0.35
Respiratory rate	0.34
Number of radiology studies ordered during the ED visit	0.34
Insurance category	0.34
Number of X-rays ordered during the ED visit	0.34
Systolic blood pressure	0.34
Weight	0.33
Chief complaint	0.32
SpO ₂ ^b	0.32
Wheezing	0.32
Retractions	0.29
Number of consults during the ED visit	0.28
Whether the patient is up-to-date with his/her immunizations	0.27
Race	0.27
Enterovirus infection	0.25
Respiratory syncytial virus infection	0.24
Coinfection	0.24
Prior hospitalization	0.22
Prior intubation	0.22
Dehydration	0.20
Language barrier to learning	0.20
Rhinovirus infection	0.20
Prematurity	0.18
History of bronchopulmonary dysplasia	0.16
History of eczema	0.15

^aED: emergency department.

^bSpO₂: peripheral capillary oxygen saturation.

Figure 2. The receiver operating characteristic curve of our model.**Table 5.** The error matrix of our predictive model.

Class	Appropriate admission	Appropriate emergency department discharge
Predicted admission	1083	241
Predicted emergency department discharge	93	2159

Table 6. A comparison of our model and the emergency department clinician's disposition decision.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC ^a	PPV ^b (%)	NPV ^c (%)
Our model	90.66	92.09	89.96	0.960	81.80	95.87
The emergency department clinician's disposition decision	93.68	98.55	91.29	0.949	84.72	99.23

^aAUC: area under the receiver operating characteristic curve.

^bPPV: positive predictive value.

^cNPV: negative predictive value.

Figure 3. The calibration plot of our model by decile of predicted probability of appropriate admission.

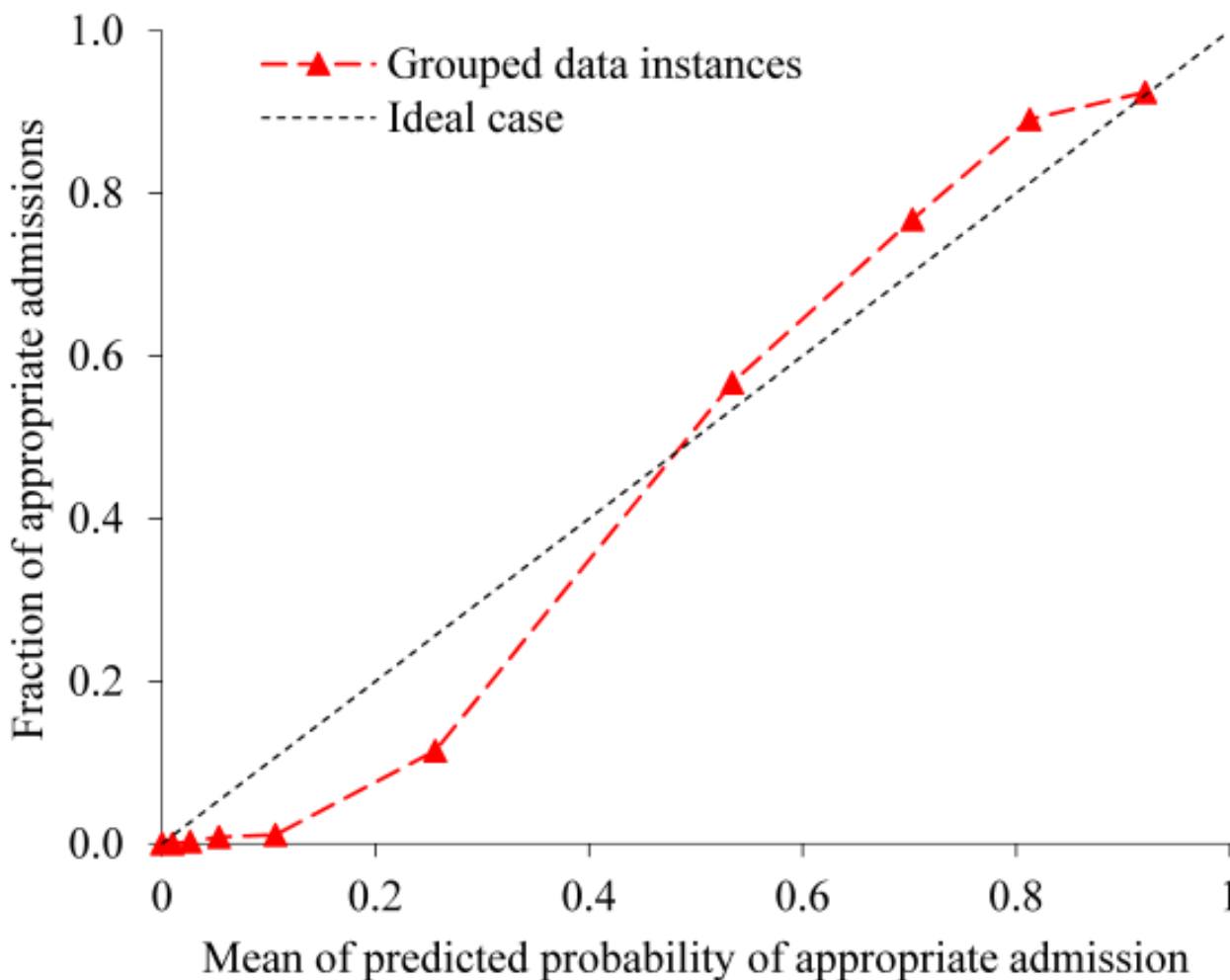


Figure 3 shows the calibration plot of our model by decile of predicted probability of appropriate admission. The Hosmer-Lemeshow test showed imperfect calibration of the predicted probabilities and the actual outcomes ($P < .001$). When the predicted probability is < 0.5 , our model tends to overestimate the actual probability. When the predicted probability is > 0.5 , our model tends to underestimate the actual probability.

Discussion

Principal Results

We developed the first machine learning classification model to accurately predict appropriate hospital admission for ED patients with bronchiolitis. Our model is a significant improvement over the previous models for predicting hospital admission in ED patients with bronchiolitis [7-9,27-29]. Our model has good accuracy, with five of the six performance metrics achieving a value $\geq 90\%$ and the other metric achieving a value $> 80\%$. Although our model attained a 3.02% lower accuracy than Intermountain Healthcare clinicians' ED disposition decisions (90.66% vs 93.68%), we still view our model as a step forward with great potential. Within 0.01 second, our model can output the prediction result for a new patient. With further improvement to boost its accuracy and automatically explain its prediction results [42,43], our model

could be integrated into an electronic health record system and become the base of a decision-support tool to help make appropriate ED disposition decisions for bronchiolitis. At that time, a clinician could use the model's output as a point of reference when considering the disposition decision. This could provide value, improve outcomes, and reduce health care costs for bronchiolitis, regardless of whether our future final model can achieve a higher accuracy than Intermountain Healthcare clinicians' ED disposition decisions. Our faith in this model stems from the following considerations:

1. Intermountain Healthcare has several collaborative partnerships among its EDs and hospitals to facilitate coordination of pediatric specialty care and has completed multiple quality-improvement projects for bronchiolitis management. About 52.16% (3963/7598) of ED visits for bronchiolitis within Intermountain Healthcare occur at a tertiary pediatric hospital with an ED staffed by pediatric-specific clinicians. On average, the ED disposition decisions for bronchiolitis made at Intermountain Healthcare could be more accurate than those made at some other healthcare systems, especially those systems with general practice physicians or fewer pediatricians working in their EDs. Our model can be valuable for those systems, if it reaches a higher accuracy than the clinicians' ED disposition decisions made at those systems. There is some

evidence indicating this possibility. Most inappropriate ED disposition decisions are unnecessary admissions [26]. In our data set, 14.36% of hospital admissions from the ED were deemed unnecessary [26]. In the literature [44,45], this percentage is reported to be larger (20%-29%). To understand our model's value for other systems, additional studies need to be conducted using data of those systems. This is an interesting area for future work.

2. **Figure 4** shows the degree of missing values of each feature with missing values. **Figure 5** shows the probability mass function of the number of features with missing values in each data instance. In our data set, several attributes have numerous missing values because those values were either recorded on paper or occasionally undocumented and therefore were not available in Intermountain Healthcare's electronic health record system. In particular, wheezing and retractions values were missing for 73.56% (5589/7598) of ED visits for bronchiolitis. Systolic and diastolic blood pressure values were missing for 46.49% (3532/7598) of ED visits for bronchiolitis. This could lower the model's accuracy. In the future, these attributes are expected to be recorded more completely in Intermountain Healthcare's newly implemented Cerner-based electronic health record system. After retraining our model on more complete Intermountain Healthcare data from future years, we would

expect its accuracy to increase. In addition, multiple other healthcare systems like Seattle Children's Hospital have been using the Cerner electronic health record system to record these attributes relatively completely for many years. Our model could possibly achieve a higher accuracy if trained with data from those systems. Both of these areas are interesting for future work.

3. When making ED disposition decisions for bronchiolitis, clinicians often face some level of uncertainty and would prefer to obtain a second opinion by a reasonably accurate predictive model, particularly if some technique is used to automatically explain the model's prediction results. For this purpose, we can use our prior method [42,43] to automatically provide rule-based explanations for any machine learning model's classification results with no accuracy loss.

When reporting the performance metrics, we used the default cut-off threshold that Weka chose in order to decide between predicted admission and predicted ED discharge. Different health care systems could emphasize different performance metrics and provide divergent weights to FPs and FNs. As is the case with predictive modeling, in general, a health care system can always adjust the cut-off threshold based on the system's preferences.

Figure 4. The degree of missing values of each feature with missing values. SpO₂: peripheral capillary oxygen saturation.

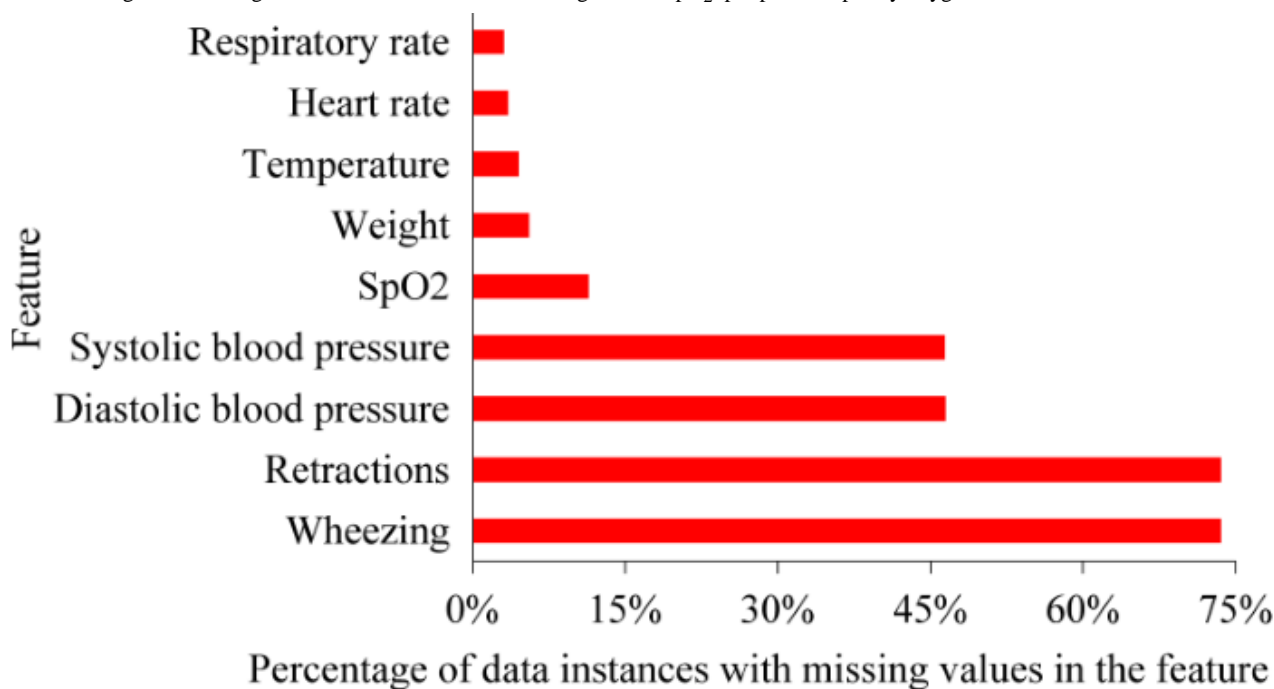
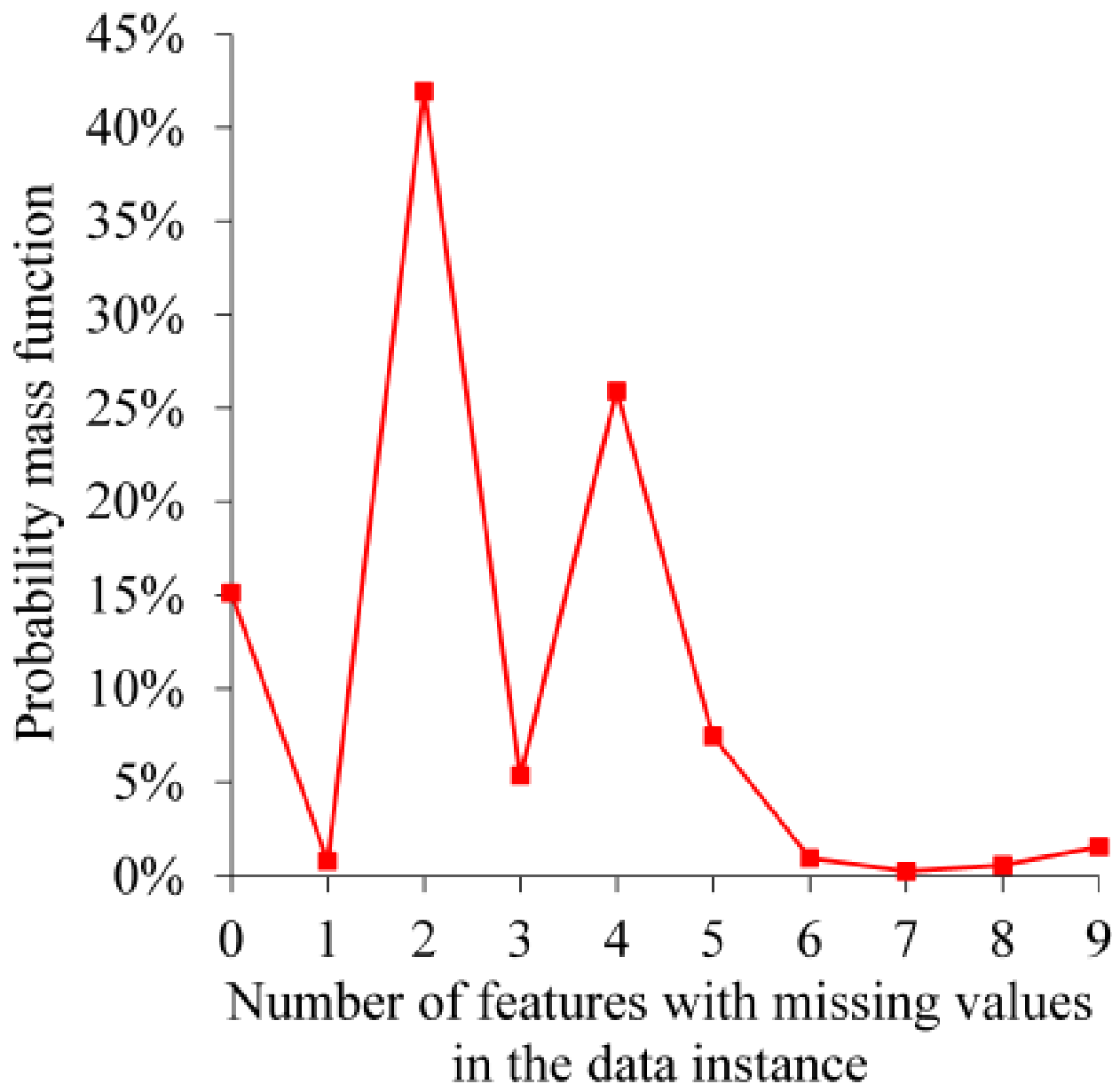


Figure 5. The probability mass function of the number of features with missing values in each data instance.



Comparison With Prior Work

Previously, researchers constructed several models to predict hospital admission in ED patients with bronchiolitis [7-9,27-29]. Table 7 compared these previous models with our model. Compared to our model, which predicts the appropriate ED disposition decision, the previous models are less accurate and incorrectly assume that actual ED disposition decisions are always appropriate. Our model uses data from more patients, more predictive features, and a more sophisticated classification algorithm than the previous models. As is the case with predictive modeling, in general, all of these features help improve our model's accuracy.

Some aspects of our findings are similar to those of previous studies. In our data set, 39.59% (3008/7598) of ED visits for bronchiolitis ended in hospitalization. This percentage is within the 32%-40% range of hospital admission rates on ED visits for bronchiolitis reported in the literature [7-9].

Limitations

This study has several limitations. First, it used data from a single health care system, Intermountain Healthcare, and did not test the generalizability of the results. In the future, studies should validate our predictive models using data from other healthcare systems. We are reasonably confident in our results, as our study was conducted in a realistic setting for finding factors generalizable to other US healthcare systems. "Intermountain Healthcare is a large healthcare system with EDs at 22 heterogeneous hospitals spread over a large geographic area, ranging from community metropolitan and rural hospitals attended by general practitioners and family doctors with constrained pediatric resources to tertiary care children's and general hospitals in urban areas attended by sub-specialists. Each hospital has a different patient population, geographic location, staff composition, scope of services, and cultural background" [26].

Table 7. A comparison of our model and several previous models for predicting emergency department disposition decisions for bronchiolitis.

Model	ED ^a visits (n)	Method for building the model	Features included in the final model	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC ^b	PPV ^c (%)	NPV ^d (%)
Our model	7599	Random forest	As listed in the Results section	90.66	92.09	89.96	0.960	81.80	95.87
Walsh et al [27]	119	Neural network ensemble	Age, respiratory rate after initial treatment, heart rate before initial treatment, oxygen saturation before and after initial treatment, dehydration, maternal smoking, increased work of breathing, poor feeding, wheezes only without associated crackles, entry temperature, and presence of both crackles and wheezes	81	78	82	— ^e	68	89
Marlais et al [7]	449	Scoring system	Age, respiratory rate, heart rate, oxygen saturation, and duration of symptoms	—	74	77	0.81	67	83
Destino et al [28]	195	Single variable	The Children's Hospital of Wisconsin respiratory score	—	65	65	0.68	—	—
Laham et al [8]	101	Logistic regression	Age, need for intravenous fluids, hypoxia, and nasal wash lactate dehydrogenase concentration	80	81	77	0.87	88	66
Corneli et al [9]	598	Decision tree	Oxygen saturation, the Respiratory Distress Assessment Instrument score computed from wheezing and retractions, and respiratory rate	—	56	74	—	—	—
Walsh et al [29]	300	Logistic regression	Age, dehydration, increased work of breathing, and heart rate	—	91	83	—	62	—

^aED: emergency department

^bAUC: area under the receiver operating characteristic curve

^cPPV: positive predictive value

^dNPV: negative predictive value

^eThe performance metric is unreported in the original paper describing the model.

Second, despite being an integrated healthcare system, Intermountain Healthcare does not have complete clinical and administrative data on all of its patients. Our data set missed information on patients' health care use that occurred at non-Intermountain Healthcare facilities. Inclusion of data from those facilities may lead to different results, but we do not expect this inclusion to significantly change our results. Intermountain Healthcare delivers ~85% of pediatric care in Utah [32]. Hence, our data set is reasonably complete with regard to capturing health care use among bronchiolitis patients in Utah.

Third, our operational definition of appropriate hospital admission is imperfect and excludes factors such as availability of patient transportation, preference of the patient's parents, and hour of ED disposition [26]. Many of these factors are often undocumented in patient records. For some hospital admissions from the ED that were regarded as unnecessary based on our

operational definition, the original admission decisions could be made because of these factors.

Finally, besides the features used in the study, other features could help improve the model's accuracy. Finding new predictive features is an interesting area for future work.

Conclusions

Our model can predict appropriate hospital admission for ED patients with bronchiolitis with good accuracy. In particular, our model achieved an AUC of 0.960. An AUC \geq 0.9 is considered outstanding discrimination [46]. With further improvement, our model could be integrated into an electronic health record system to provide personalized real-time decision support for making ED disposition decisions for bronchiolitis, which could help standardize care and improve outcomes for bronchiolitis.

Acknowledgments

We thank Farrant Sakaguchi, Michael Mundorff, Karen Valentine, Chris Benitez, JoAnn Banks, Bart Dodds, Xiaoming Sheng, and Jim Bradshaw for their helpful discussions and help in retrieving the Intermountain Healthcare data set. GL, BLS, MDJ, FLN, and SH were partially supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under

Award Number R21HL128875. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

GL was mainly responsible for the paper. He conceptualized and designed the study, performed the literature review and data analysis, and wrote the paper. BLS, MDJ, and FLN provided feedback on various medical issues, contributed to conceptualizing the presentation, and revised the paper. SH took part in retrieving the Intermountain Healthcare data set and interpreting its detected peculiarities.

Conflicts of Interest

None declared.

References

1. Zorc JJ, Hall CB. Bronchiolitis: recent evidence on diagnosis and management. *Pediatrics* 2010 Feb;125(2):342-349. [doi: [10.1542/peds.2009-2092](https://doi.org/10.1542/peds.2009-2092)] [Medline: [20100768](https://pubmed.ncbi.nlm.nih.gov/20100768/)]
2. Hasegawa K, Tsugawa Y, Brown DFM, Mansbach JM, Camargo CA. Trends in bronchiolitis hospitalizations in the United States, 2000-2009. *Pediatrics* 2013 Jul;132(1):28-36 [FREE Full text] [doi: [10.1542/peds.2012-3877](https://doi.org/10.1542/peds.2012-3877)] [Medline: [23733801](https://pubmed.ncbi.nlm.nih.gov/23733801/)]
3. Mansbach JM, Emond JA, Camargo CA. Bronchiolitis in US emergency departments 1992 to 2000: epidemiology and practice variation. *Pediatr Emerg Care* 2005 Apr;21(4):242-247. [Medline: [15824683](https://pubmed.ncbi.nlm.nih.gov/15824683/)]
4. Parker MJ, Allen U, Stephens D, Lalani A, Schuh S. Predictors of major intervention in infants with bronchiolitis. *Pediatr Pulmonol* 2009 Apr;44(4):358-363. [doi: [10.1002/ppul.21010](https://doi.org/10.1002/ppul.21010)] [Medline: [19283838](https://pubmed.ncbi.nlm.nih.gov/19283838/)]
5. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, Anderson LJ. Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA* 1999 Oct 20;282(15):1440-1446. [Medline: [10535434](https://pubmed.ncbi.nlm.nih.gov/10535434/)]
6. Hasegawa K, Tsugawa Y, Brown DFM, Mansbach JM, Camargo CA. Temporal trends in emergency department visits for bronchiolitis in the United States, 2006 to 2010. *Pediatr Infect Dis J* 2014 Jan;33(1):11-18 [FREE Full text] [doi: [10.1097/INF.0b013e3182a5f324](https://doi.org/10.1097/INF.0b013e3182a5f324)] [Medline: [23934206](https://pubmed.ncbi.nlm.nih.gov/23934206/)]
7. Marlais M, Evans J, Abrahamson E. Clinical predictors of admission in infants with acute bronchiolitis. *Arch Dis Child* 2011 Jul;96(7):648-652. [doi: [10.1136/adc.2010.201079](https://doi.org/10.1136/adc.2010.201079)] [Medline: [21339199](https://pubmed.ncbi.nlm.nih.gov/21339199/)]
8. Laham FR, Trott AA, Bennett BL, Kozinetz CA, Jewell AM, Garofalo RP, et al. LDH concentration in nasal-wash fluid as a biochemical predictor of bronchiolitis severity. *Pediatrics* 2010 Feb;125(2):e225-e233 [FREE Full text] [doi: [10.1542/peds.2009-0411](https://doi.org/10.1542/peds.2009-0411)] [Medline: [20100751](https://pubmed.ncbi.nlm.nih.gov/20100751/)]
9. Corneli HM, Zorc JJ, Holubkov R, Bregstein JS, Brown KM, Mahajan P, Bronchiolitis Study Group for the Pediatric Emergency Care Applied Research Network. Bronchiolitis: clinical characteristics associated with hospitalization and length of stay. *Pediatr Emerg Care* 2012 Feb;28(2):99-103. [doi: [10.1097/PEC.0b013e3182440b9b](https://doi.org/10.1097/PEC.0b013e3182440b9b)] [Medline: [22270499](https://pubmed.ncbi.nlm.nih.gov/22270499/)]
10. Scottish IGN. Guideline Central. Bronchiolitis in children. A national clinical guideline URL: <https://www.guidelinecentral.com/summaries/bronchiolitis-in-children-a-national-clinical-guideline/> [accessed 2019-01-17] [WebCite Cache ID 6bhQvWHZA]
11. AHRQ - Agency for Healthcare Research and Quality: Advancing Excellence in Health Care. Cincinnati Children's Hospital Medical Center: Evidence-based care guideline for management of first time episode bronchiolitis in infants less than 1 year of age URL: <http://www.guideline.gov/content.aspx?id=34411> [accessed 2019-01-17] [WebCite Cache ID 6bhOyjITC]
12. Brand PL, Vaessen-Verberne AA. Differences in management of bronchiolitis between hospitals in The Netherlands. *Dutch Paediatric Respiratory Society. Eur J Pediatr* 2000 May;159(5):343-347. [Medline: [10834519](https://pubmed.ncbi.nlm.nih.gov/10834519/)]
13. Ducharme FM. Management of acute bronchiolitis. *BMJ* 2011 Apr 06;342:d1658. [doi: [10.1136/bmj.d1658](https://doi.org/10.1136/bmj.d1658)] [Medline: [21471173](https://pubmed.ncbi.nlm.nih.gov/21471173/)]
14. Behrendt CE, Decker MD, Burch DJ, Watson PH. International variation in the management of infants hospitalized with respiratory syncytial virus. International RSV Study Group. *Eur J Pediatr* 1998 Mar;157(3):215-220. [Medline: [9537488](https://pubmed.ncbi.nlm.nih.gov/9537488/)]
15. Chamberlain JM, Patel KM, Pollack MM. Association of emergency department care factors with admission and discharge decisions for pediatric patients. *J Pediatr* 2006 Nov;149(5):644-649. [doi: [10.1016/j.jpeds.2006.05.047](https://doi.org/10.1016/j.jpeds.2006.05.047)] [Medline: [17095336](https://pubmed.ncbi.nlm.nih.gov/17095336/)]
16. Johnson DW, Adair C, Brant R, Holmwood J, Mitchell I. Differences in admission rates of children with bronchiolitis by pediatric and general emergency departments. *Pediatrics* 2002 Oct;110(4):e49. [Medline: [12359822](https://pubmed.ncbi.nlm.nih.gov/12359822/)]
17. Mallory MD, Shay DK, Garrett J, Bordley WC. Bronchiolitis management preferences and the influence of pulse oximetry and respiratory rate on the decision to admit. *Pediatrics* 2003 Jan;111(1):e45-e51. [Medline: [12509594](https://pubmed.ncbi.nlm.nih.gov/12509594/)]
18. Plint AC, Johnson DW, Wiebe N, Bulloch B, Pusic M, Joubert G, et al. Practice variation among pediatric emergency departments in the treatment of bronchiolitis. *Acad Emerg Med* 2004 Apr;11(4):353-360 [FREE Full text] [Medline: [15064208](https://pubmed.ncbi.nlm.nih.gov/15064208/)]
19. Vogel AM, Lennon DR, Harding JE, Pinnock RE, Graham DA, Grimwood K, et al. Variations in bronchiolitis management between five New Zealand hospitals: can we do better? *J Paediatr Child Health* 2003;39(1):40-45. [Medline: [12542811](https://pubmed.ncbi.nlm.nih.gov/12542811/)]

20. Willson DF, Horn SD, Hendley JO, Smout R, Gassaway J. Effect of practice variation on resource utilization in infants hospitalized for viral lower respiratory illness. *Pediatrics* 2001 Oct;108(4):851-855. [Medline: [11581435](#)]
21. Willson DF, Jiao JH, Hendley JO, Donowitz L. Invasive monitoring in infants with respiratory syncytial virus infection. *J Pediatr* 1996 Mar;128(3):357-362. [Medline: [8774504](#)]
22. Wang EE, Law BJ, Boucher FD, Stephens D, Robinson JL, Dobson S, et al. Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) study of admission and management variation in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. *J Pediatr* 1996 Sep;129(3):390-395. [Medline: [8804328](#)]
23. Christakis DA, Cowan CA, Garrison MM, Molteni R, Marcuse E, Zerr DM. Variation in inpatient diagnostic testing and management of bronchiolitis. *Pediatrics* 2005 Apr;115(4):878-884. [doi: [10.1542/peds.2004-1299](#)] [Medline: [15805359](#)]
24. Mansbach JM, Clark S, Christopher NC, LoVecchio F, Kunz S, Acholonu U, et al. Prospective multicenter study of bronchiolitis: predicting safe discharges from the emergency department. *Pediatrics* 2008 Apr;121(4):680-688. [doi: [10.1542/peds.2007-1418](#)] [Medline: [18381531](#)]
25. McBride SC, Chiang VW, Goldmann DA, Landrigan CP. Preventable adverse events in infants hospitalized with bronchiolitis. *Pediatrics* 2005 Sep;116(3):603-608. [doi: [10.1542/peds.2004-2387](#)] [Medline: [16140699](#)]
26. Luo G, Johnson MD, Nkoy FL, He S, Stone BL. Appropriateness of Hospital Admission for Emergency Department Patients with Bronchiolitis: Secondary Analysis. *JMIR Med Inform* 2018 Nov 05;6(4):e10498 [FREE Full text] [doi: [10.2196/10498](#)] [Medline: [30401659](#)]
27. Walsh P, Cunningham P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. An artificial neural network ensemble to predict disposition and length of stay in children presenting with bronchiolitis. *Eur J Emerg Med* 2004 Oct;11(5):259-264. [Medline: [15359198](#)]
28. Destino L, Weisgerber MC, Soung P, Bakalarski D, Yan K, Rehborg R, et al. Validity of respiratory scores in bronchiolitis. *Hosp Pediatr* 2012 Oct;2(4):202-209 [FREE Full text] [Medline: [24313026](#)]
29. Walsh P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. A validated clinical model to predict the need for admission and length of stay in children with acute bronchiolitis. *Eur J Emerg Med* 2004 Oct;11(5):265-272. [Medline: [15359199](#)]
30. Luo G, Nkoy FL, Gesteland PH, Glasgow TS, Stone BL. A systematic review of predictive modeling for bronchiolitis. *Int J Med Inform* 2014 Oct;83(10):691-714. [doi: [10.1016/j.ijmedinf.2014.07.005](#)] [Medline: [25106933](#)]
31. Luo G, Stone BL, Johnson MD, Nkoy FL. Predicting Appropriate Admission of Bronchiolitis Patients in the Emergency Department: Rationale and Methods. *JMIR Res Protoc* 2016 Mar 07;5(1):e41 [FREE Full text] [doi: [10.2196/resprot.5155](#)] [Medline: [26952700](#)]
32. Byington CL, Reynolds CC, Korgenski K, Sheng X, Valentine KJ, Nelson RE, et al. Costs and infant outcomes after implementation of a care process model for febrile infants. *Pediatrics* 2012 Jul;130(1):e16-e24 [FREE Full text] [doi: [10.1542/peds.2012-0127](#)] [Medline: [22732178](#)]
33. Flaherman VJ, Ragins AI, Li SX, Kipnis P, Masaquel A, Escobar GJ. Frequency, duration and predictors of bronchiolitis episodes of care among infants ≥ 32 weeks gestation in a large integrated healthcare system: a retrospective cohort study. *BMC Health Serv Res* 2012 Jun 08;12:144 [FREE Full text] [doi: [10.1186/1472-6963-12-144](#)] [Medline: [22682080](#)]
34. Mittal V, Darnell C, Walsh B, Mehta A, Badawy M, Morse R, et al. Inpatient bronchiolitis guideline implementation and resource utilization. *Pediatrics* 2014 Mar;133(3):e730-e737 [FREE Full text] [doi: [10.1542/peds.2013-2881](#)] [Medline: [24534398](#)]
35. Sandweiss DR, Mundorff MB, Hill T, Wolfe D, Greene T, Andrews S, et al. Decreasing hospital length of stay for bronchiolitis by using an observation unit and home oxygen therapy. *JAMA Pediatr* 2013 May;167(5):422-428. [doi: [10.1001/jamapediatrics.2013.1435](#)] [Medline: [23479000](#)]
36. Steyerberg EW. *Clinical Prediction Models: A Practical Approach To Development, Validation, And Updating (statistics For Biology And Health)*. New York, NY: Springer-Verlag; 2009.
37. Witten IH, Frank E, Hall MA, Pal CJ. *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, MA: Morgan Kaufmann; 2016.
38. Kaggle. URL: <https://www.kaggle.com/> [accessed 2019-01-17] [WebCite Cache ID 6oOi4j9t]
39. Zeng X, Luo G. Progressive sampling-based Bayesian optimization for efficient and automatic machine learning model selection. *Health Inf Sci Syst* 2017 Dec;5(1):2 [FREE Full text] [doi: [10.1007/s13755-017-0023-z](#)] [Medline: [29038732](#)]
40. Feudtner C, Feinstein JA, Zhong W, Hall M, Dai D. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr* 2014 Aug 08;14:199 [FREE Full text] [doi: [10.1186/1471-2431-14-199](#)] [Medline: [25102958](#)]
41. Agresti A. *Categorical Data Analysis*. Hoboken, NJ: Wiley; 2012.
42. Luo G. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. *Health Inf Sci Syst* 2016;4:2 [FREE Full text] [doi: [10.1186/s13755-016-0015-4](#)] [Medline: [26958341](#)]
43. Luo G. A roadmap for semi-automatically extracting predictive and clinically meaningful temporal features from medical data for predictive modeling. *Global Transitions* 2019;1(1) (forthcoming).
44. Shaw KN, Bell LM, Sherman NH. Outpatient assessment of infants with bronchiolitis. *Am J Dis Child* 1991 Feb;145(2):151-155. [Medline: [1994678](#)]

45. Shafik MH, Seoudi TMM, Raway TS, Al Harbash NZ, Ahmad MMA, Al Mutairi HF. Appropriateness of pediatric hospitalization in a general hospital in Kuwait. *Med Princ Pract* 2012;21(6):516-521 [FREE Full text] [doi: [10.1159/000339084](https://doi.org/10.1159/000339084)] [Medline: [22678120](https://pubmed.ncbi.nlm.nih.gov/22678120/)]
46. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. Hoboken, NJ: Wiley; 2013.

Abbreviations

AUC: area under the receiver operating characteristic curve

ED: emergency department

FN: false negative

FP: false positive

ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification

NPV: negative predictive value

PPV: positive predictive value

SpO₂: peripheral capillary oxygen saturation

TN: true negative

TP: true positive

Edited by G Eysenbach; submitted 23.10.18; peer-reviewed by K Chen, P Giabbanelli, J op den Buijs, R Rivas; comments to author 21.11.18; revised version received 27.11.18; accepted 12.12.18; published 22.01.19.

Please cite as:

Luo G, Stone BL, Nkoy FL, He S, Johnson MD

Predicting Appropriate Hospital Admission of Emergency Department Patients with Bronchiolitis: Secondary Analysis

JMIR Med Inform 2019;7(1):e12591

URL: <http://medinform.jmir.org/2019/1/e12591/>

doi: [10.2196/12591](https://doi.org/10.2196/12591)

PMID: [30668518](https://pubmed.ncbi.nlm.nih.gov/30668518/)

©Gang Luo, Bryan L Stone, Flory L Nkoy, Shan He, Michael D Johnson. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 22.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Development of Prediction Models Using Machine Learning Algorithms for Girls with Suspected Central Precocious Puberty: Retrospective Study

Liyan Pan^{1*}, MSc; Guangjian Liu^{1*}, PhD; Xiaojian Mao^{2*}, PhD; Huixian Li¹, MSc; Jiexin Zhang¹, BD; Huiying Liang^{1*}, PhD; Xiuzhen Li^{2*}, PhD

¹Institute of Pediatrics, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

²Department of Genetics and Endocrinology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

*these authors contributed equally

Corresponding Author:

Xiuzhen Li, PhD

Department of Genetics and Endocrinology

Guangzhou Women and Children's Medical Center

Guangzhou Medical University

9 Jinsui Road

Tianhe District

Guangzhou, 510623

China

Phone: 86 02038857692

Fax: 86 02038857692

Email: 13725100840@163.com

Abstract

Background: Central precocious puberty (CPP) in girls seriously affects their physical and mental development in childhood. The method of diagnosis—gonadotropin-releasing hormone (GnRH)—stimulation test or GnRH analogue (GnRHa)—stimulation test—is expensive and makes patients uncomfortable due to the need for repeated blood sampling.

Objective: We aimed to combine multiple CPP-related features and construct machine learning models to predict response to the GnRHa-stimulation test.

Methods: In this retrospective study, we analyzed clinical and laboratory data of 1757 girls who underwent a GnRHa test in order to develop XGBoost and random forest classifiers for prediction of response to the GnRHa test. The local interpretable model-agnostic explanations (LIME) algorithm was used with the black-box classifiers to increase their interpretability. We measured sensitivity, specificity, and area under receiver operating characteristic (AUC) of the models.

Results: Both the XGBoost and random forest models achieved good performance in distinguishing between positive and negative responses, with the AUC ranging from 0.88 to 0.90, sensitivity ranging from 77.91% to 77.94%, and specificity ranging from 84.32% to 87.66%. Basal serum luteinizing hormone, follicle-stimulating hormone, and insulin-like growth factor-I levels were found to be the three most important factors. In the interpretable models of LIME, the abovementioned variables made high contributions to the prediction probability.

Conclusions: The prediction models we developed can help diagnose CPP and may be used as a prescreening tool before the GnRHa-stimulation test.

(*JMIR Med Inform* 2019;7(1):e11728) doi:[10.2196/11728](https://doi.org/10.2196/11728)

KEYWORDS

central precocious puberty; GnRHa-stimulation test; machine learning; prediction model

Introduction

Precocious puberty is related to the development of secondary sexual characteristics in girls before the age of 8 years and in boys before the age of 9 years. In recent years, the age of puberty onset has shown a decreasing trend, and puberty is related to subsequent health outcomes such as breast cancer, diabetes, and behavioral disorders [1]. Central precocious puberty (CPP), also known as true precocious puberty, is caused by early activation of the hypothalamic-pituitary-gonadal axis with clinical pubertal symptoms. CPP can influence final adult height and result in psychological problems, which will cause inappropriate behaviors. It is important for girls with suspected CPP to be evaluated and diagnosed in a timely manner.

The gold standard in the confirmation of CPP is the positive response of gonadotropin to a gonadotropin-releasing hormone (GnRH)-stimulation test. In the absence of GnRH, GnRH analogues (GnRHa) are usually used instead [2]. However, the stimulation test is time consuming and expensive. Besides, the test is painful and make patients uncomfortable due to the need for repeated blood sampling at different time points. Therefore, another method to avoid the disadvantages of the GnRHa-stimulation test will be of great help in the diagnosis of CPP.

Several studies have focused on investigating a single gonadotropin biomarker to identify patients with CPP conveniently. Basal or peak serum luteinizing hormone (LH), follicle-stimulation hormone (FSH), and the ratio LH/FSH are the most common biomarkers reported [3-7]. However, the cut-off values of these single biomarkers depend on the assay used to measure the gonadotropin levels. As a result, cut-off points in previous studies differed widely. Moreover, both the Pasternak group [3] and the Mogensen group [8] reported that a single basal serum LH measurement could verify the presence of CPP, but could not confirm the absence of CPP. Therefore, a single gonadotropin parameter alone may not be sufficient for the diagnosis of CPP, and clinical and laboratory factors that can predict response to the GnRHa-stimulation test should be combined [9,10]. Suh et al [10] found that accelerated growth rate, advanced bone age, and increased basal gonadotropin and insulin-like growth factor-I (IGF-I) levels were correlated with CPP. Traditional statistical analysis including *t* test and binary logistic regression were used to select factors correlated with the GnRH test [11-14]. Although remarkable progress has been made in these studies, there is a long way to go for their application in clinics due to the low sensitivity or specificity of tests.

Considering the previous studies and the extensive application of machine learning algorithms in the medical field, we aimed to determine whether combining multiple variables with machine learning classifiers could improve prediction of the GnRHa-stimulation test and thus help diagnose CPP.

Methods

Population and Variables

We enrolled 1757 girls with CPP symptoms who visited the Pediatric Day Ward of the Endocrinology Department at Guangzhou Women and Children's Medical Center from January 2012 to March 2018. All subjects had undergone the GnRHa-stimulation test. Girls with any other disorders or intracranial lesions were excluded from the study.

Girls fulfilling the following eligibility criteria were considered to have a positive response to the GnRHa-stimulation test and were diagnosed with CPP in our study: (1) peak LH level ≥ 10 IU/L or peak LH level ≥ 5 IU/L combined with a ratio of peak LH to FSH value ≥ 0.6 and (2) onset of secondary sexual characteristics at the age < 8 years. Girls whose laboratory tests did not satisfy all the abovementioned criteria were considered to have a negative response. According to the long-term clinical practices, the first condition with a peak LH ≥ 10 IU/L is used as the diagnosing criterion in our hospital. Peak LH level ≥ 5 IU/L combined with a ratio of peak LH to FSH value ≥ 0.6 is widely used in China and some other countries for children undergoing the GnRH-stimulation test [15,16]. Since the stimulation effect of GnRHa is almost hundreds times that of GnRH [17], a condition that affects the levels of sex hormones due to GnRH would do the same with GnRHa. Our diagnostic criteria are an improved version of the existing criteria that are adapted to our population.

Information such as chief complaints, development of secondary sexual characters, and abnormal duration of puberty were stored as free text in the clinical records of the electronic medical records system. Laboratory test values were reported as structured data in the laboratory information system. In total, 19 variables were extracted from the clinical records and laboratory results for all the 1757 patients. Specifically, 10 variables extracted from the clinical records were weight, height, body mass index (BMI), abnormal duration of puberty in records (History), menarche, core in breast (Core), pigmentation, development stage of pubes (Pubes), development stage of left breast, and development stage of right breast. Breast and pubic hair development were evaluated using Tanner staging (stages 1 to 5). Nine variables extracted from the laboratory results were age, basal serum LH, FSH, estradiol, prolactin, testosterone, growth hormone, IGF-I, IGF-binding protein-3 levels before the GnRHa test.

Among the 1757 patients, 436 girls had examination reports available, including pelvic ultrasonography (for development of the uterus and ovaries) and radiography of the left hand (for bone age). Six variables extracted from the examination reports were development of uterus, existence of follicle, uterine volume, left and right ovary volumes, and bone age. Bone age was measured by the Greulich and Pyle method [18]. The variables from the clinical records and the examination reports were extracted first with traditional regex match using Python [19] and then examined manually by two endocrinologists.

This study was approved by the Institutional Review Board of Guangzhou Women and Children's Medical Center and

conducted in accordance with the ethical guidelines of the Declaration of Helsinki of the World Medical Association. The requirement to obtain informed consent was waived because of the retrospective nature of the study. Data used in this study were anonymous, and no identifiable personal data of the patients were available for the analysis.

Data Preprocessing

Variables with more than 20% missing data, such as growth rate of height and weight and heights of parents, were excluded from this study. Missing values for continuous variables were replaced with mean values of all the samples grouped by age. Discrete variables like experience of menarche were filled with a value of 0. Discrete variables like Tanner stage for breast and pubes were filled with the least degree (stage 1).

Model Development and Assessment

Tree learning classifiers allow nonlinear interactions between features and have good interpretability. Considering this, we selected two tree-based ensemble binary classification algorithms—extreme gradient boosting (XGBoost) and random forest—to develop our models. We also used linear supported vector machines (SVM) and decision trees for the classification to compare the performance between ensemble models and nonensemble models. The models aimed to identify relationships between the input features and the output GnRHa test results, thereby determining whether a patient responds positively to the GnRHa test.

XGBoost is a scalable tree boosting and effective learning algorithm [20]. It trains a sequence of models to minimize errors made by existing models. Models in XGBoost are decision trees. Many data scientists have applied this algorithm to solve classification problems and achieved excellent results. XGBoost has also been successfully used in medical studies [21,22]. As XGBoost is essentially a gradient boosting tree model, which is not based on distance, normalization is not required. Random forest is another classical ensemble learning algorithm with a combination of a large amount of trees [23], which trains decision trees in parallel by using data with replacement. It applies bootstrap aggregating to tree learners, which leads to better performance as variance decreases. Random forest has the ability to handle nonlinear data and is robust to noise. Besides, parameter tuning is not that complex for these two algorithms compared to other ensemble learning algorithms. SVM is a binary classifier with a maximum margin hyperplane [24]. The decision tree classifier is a tree-like model used for classification [25].

In order to obtain robust assessments and prevent overfitting, we used a nested cross-validation with an outer Monte Carlo cross-validation [26] (MCCV, repeated 20 times) and an inner k-fold cross-validation (k=5) for parameter tuning, yielding a total of 20 times the five-fold cross-validation. In the outer MCCV loop, the whole data set is randomly divided into the training set (80%) and the test set (20%) for 20 times. For each training set, the inner stratified five-fold cross-validation loop is performed as follows. The training set is split into five subsets, where four subsets are used for training and one is used for test. Parameter tuning is performed with grid search in the inner

cross-validation. Finally, a model fitted on the training set with parameters that has the best area under the curve (AUC) evaluated on the inner test set is determined. The detailed training and test process is presented in Figure 1.

Feature Importance

XGBoost and random forest classifier have the ability to evaluate the importance of features. Feature importance is a feature weight and can represent the contribution to prediction. In XGBoost, feature importance is computed by the sum of times that the feature is selected as a tree node. In random forest, feature importance is calculated based on the out-of-bag (OOB) error. OOB error is the mean prediction error for training observations in the respective bootstrap sample. After randomly adding noise perturbations to OOB samples, a feature with a higher OOB error difference is more important, with higher feature importance. For both models, each feature obtained 20 feature importance values with 20 times the MCCV. We summed all the feature importance values for each feature and obtained a rank for all features.

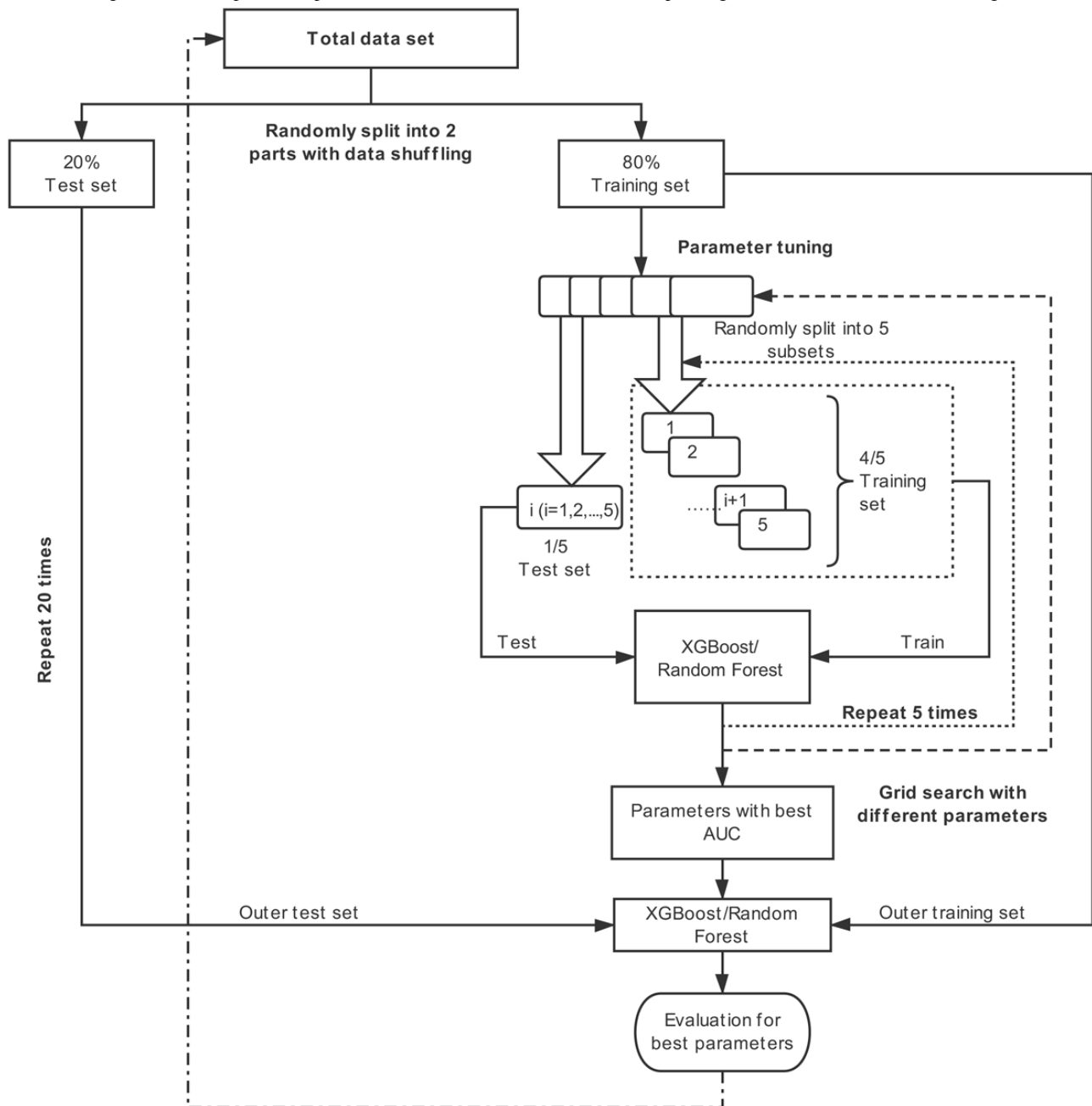
Model Interpretation

One disadvantage of machine learning is that the model usually runs as a black box. However, it is necessary for a doctor to understand the reasons why a model makes such a prediction in the clinic, especially when timely detection is necessary. Tree-based models can provide feature importance at a global level but not in a specific case. The local interpretable model-agnostic explanations (LIME) algorithm is developed to identify an interpretable model that is locally faithful for each individual prediction [27,28]. It provides relative feature contributions for a single instance of the prediction result. LIME generates neighborhood data by randomly perturbing features from the instance. It then learns locally weighted linear models on this neighborhood data to explain each of the classes in an interpretable way. Parameters in LIME mainly include the maximum number of features in explanation, number of neighborhood samples to generate, and machine learning prediction function. We used the LIME library from the original authors for the model interpretation. The number of neighborhood samples is 5000 by default. The parameter num_features (maximum number of features) is the number of features shown in the explanation. The default value is 10. In our study, the value was 9 for a clear layout, as the contributions of features ranked after 9 were almost zero.

In our study, we used a submodular pick [27] instead of a random pick to select a diverse, representative set of samples from the test set for nonredundant explanations. For these samples, we then obtained the class probabilities, and the representative individuals were assigned with an average weight contributed by each feature to display how the classifier made a decision. Finally, we went over all these LIME results together with the endocrinologists to decide whether we should trust the results of the model.

All computation and visualization were performed in Python [19] using packages like Scikit-learn, Pandas, Lime, and Matplotlib.

Figure 1. Training and validation process of prediction models. AUC: area under receiver operating characteristic; XGBoost: extreme gradient boosting.



Results

Subject Characteristics

Among the 1757 girls included in our study, 966 were positive for the GnRHa-stimulation test and diagnosed with CPP; the remaining 791 girls showed a negative response to the test. As shown in Table 1, 16 of the 19 variables were significantly higher in the CPP group than in the non-CPP group ($P < .05$), whereas prolactin, BMI, and pigmentation were similar in both groups.

Evaluation for Models

First, we developed prediction models with the data of 19 clinical and laboratory variables (Table 1) from all 1757 patients.

Two machine learning algorithms, XGBoost and random forest classifiers, were used, and parameters with the best AUC were selected for each model. The performance as well as the selected parameters of the models are listed in Table 2, and their receiver operating characteristic (ROC) curves are plotted in Figure 2. The performance was evaluated on the 20 test sets split with MCCV. Both models had strong prediction powers, with a specificity of $\geq 84.32\%$, a sensitivity of $\geq 77.91\%$, and an AUC of ≥ 0.88 . The XGBoost classifier is slightly more effective than the random forest classifier, especially in terms of the specificity ($P < .01$), whereas random forest is much more efficient in terms of the computation speed with less model complexity.

Table 1. Basic characteristics of girls who underwent the GnRHa-stimulation test.

Variables	Non-CPP ^a (n=966), mean (SD)	CPP (n=791), mean (SD)	<i>P</i> value ^b
Age (years)	7.07 (1.11)	7.52 (0.99)	<.001
LH ^c (IU/L)	0.12 (0.23)	0.93 (1.28)	<.001
FSH ^d (IU/L)	1.82 (1.30)	3.01 (1.62)	<.001
GH ^e (ng/mL)	3.27 (3.26)	4.75 (4.69)	<.001
IGF-I ^f (ng/mL)	231.35 (65.93)	317.87 (89.84)	<.001
IGFBP-3 ^g (µg/mL)	4.55 (0.52)	4.81 (0.55)	<.001
Estradiol (pmol/L)	102.56 (50.96)	125.81 (60.97)	<.001
Prolactin (ng/mL)	8.73 (5.39)	8.59 (5.61)	.52
Testosterone (nmol/L)	0.80 (0.39)	0.94 (0.49)	<.001
History ^h (months)	7.67 (10.39)	9.27 (9.63)	<.001
Menstruation/menarche (yes, no)	N/A ⁱ	N/A	.03
Height ^j (cm)	127.16 (8.61)	131.61 (8.42)	<.001
Weight ^j (kg)	27.32 (5.32)	29.60 (4.95)	<.001
BMI ^k (kg/m ²)	16.73 (2.30)	16.91 (1.96)	.34
Breast core (yes, no)	N/A	N/A	.02
Pubes ^l (1-5)	1.06 (0.27)	1.14 (0.44)	<.001
Pigmentation (yes, no)	N/A	N/A	.87
Left breast ^l (1-5)	2.33 (0.84)	2.76 (0.92)	<.001
Right breast ^l (1-5)	2.32 (0.84)	2.78 (0.92)	<.001

^aCPP: central precocious puberty.

^bThe equality of each indicator was evaluated by Chi-square or Student *t* test. *P*<.05 was considered significant.

^cLH: luteinizing hormone.

^dFSH: follicle-stimulation hormone.

^eGH: growth hormone.

^fIGF-I: insulin-like growth factor-I.

^gIGFBP-3: insulin-like growth factor binding protein-3.

^hAbnormal duration in records.

ⁱN/A: not applicable.

^jAt stimulation test.

^kBMI: body mass index.

^lTanner stage.

Table 2. Predictive performance of classifiers and the corresponding parameters. A paired *t* test was performed on specificity and sensitivity for comparison against XGBoost.

Algorithms/Variables	Specificity ^a (%), mean (SD)	Sensitivity ^b (%), mean (SD)	AUC ^c , mean (SD)	Parameters
19 variables, 1757 patients				
XGBoost ^d	85.39 (1.38)	77.94 (3.50)	0.89 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
Random forest	84.32 (1.88) ^e	77.91 (3.59) ^f	0.88 (0.02)	Max depth=3, criterion=gini, number of trees=20
SVM ^g	88.94 (1.76) ^e	62.36 (4.12) ^e	0.86 (0.04)	Kernel=linear, penalty coefficient=5
Decision tree	75.90 (2.47) ^e	71.71 (3.99) ^e	0.74 (0.02)	Criterion=entropy
19 variables, 436 patients				
XGBoost	83.17 (5.29)	75.28 (6.43)	0.86 (0.04)	Learning rate=0.01, max depth=3, number of trees=500
Random forest	83.46 (6.28) ^f	74.72 (6.43) ^f	0.85 (0.04)	Max depth=3, criterion=gini, number of trees=20
SVM	88.94 (4.90) ^e	62.36 (7.73) ^e	0.86 (0.02)	Kernel=linear, penalty coefficient=5
Decision tree	76.25 (7.07) ^e	68.06 (7.12) ^e	0.72 (0.04)	Criterion=entropy
25 variables, 436 patients				
XGBoost ^d	87.66 (5.52)	76.64 (6.51)	0.90 (0.04)	Learning rate=0.01, max depth=4, number of trees=500
Random forest	87.41 (4.22) ^f	75.03 (7.91) ^f	0.90 (0.05)	Max depth=3, criterion=entropy, number of trees=20
SVM	89.81 (4.28) ^f	66.53 (7.01) ^e	0.86 (0.02)	Kernel=linear, penalty coefficient=5
Decision tree	76.35 (5.51) ^e	68.61 (7.16) ^e	0.72 (0.05)	Criterion=entropy
1-3 variables, 1757 patients, XGBoost^d				
LH ^h , IGF-I ⁱ , FSH ^j	83.17 (1.62)	76.39 (3.57)	0.86 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
LH ^h , IGF-I ⁱ	83.27 (1.62)	75.69 (3.61)	0.86 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
LH ^h , FSH ^j	83.56 (1.94)	75.83 (3.13)	0.84 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
LH ^h	83.37 (2.00)	75.97 (3.74)	0.84 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
IGF-I ⁱ , FSH ^j	80.77 (2.47)	57.08 (3.29)	0.77 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
IGF-I ⁱ	80.19 (3.14)	53.19 (4.55)	0.73 (0.02)	Learning rate=0.01, max depth=3, number of trees=500
FSH ^j	84.13 (3.87)	45.00 (5.34)	0.68 (0.02)	Learning rate=0.01, max depth=3, number of trees=500

^aSpecificity=number of true negatives/(number of true negatives+number of false positives).

^bSensitivity=number of true positives/(number of true positives+number of false negatives).

^cAUC, area under the receiver operating curve.

^dXGBoost: extreme gradient boosting.

^e*P*<.01

^fNot significant.

^gSVM: supported vector machines.

^hLH: luteinizing hormone.

ⁱIGF-I: insulin-like growth factor-I.

^jFSH: follicle-stimulation hormone.

In the data set, 436 girls had additional examination reports, and we extracted six variables from these reports (see Population and Variables subsection). To investigate whether adding image features could enhance the prediction efficiency, we combined the six variables with the 19 variables and trained and evaluated both the XGBoost and random forest models on the 436 samples, of which 180 patients belonged to the CPP group and 256

belonged to the non-CPP group. For the ease of comparison, we retrained the previous 19-variable models with the 436 samples. As shown in Table 2, the reduction in sample size led to a serious decline in model performance, whereas the addition of six image features improved their performance. Specifically, for XGBoost in 436 samples, the specificity increased from 83.17% for 19 variables to 87.66% for 25 variables, the

sensitivity increased from 75.28% to 76.64%, and the AUC increased from 0.86 to 0.90. For random forest in 436 samples with 25 variables, the specificity increased from 83.46% to 87.41%, the sensitivity increased from 74.72% to 75.03%, and the AUC increased from 0.86 to 0.90 compared to the results from 436 samples with 19 variables. Similarly, as seen in the ROC curves shown in Figure 2, XGBoost performed slightly better than the random forest classifier.

To compare performance between ensemble models and nonensemble models, the SVM and decision tree classifiers were used to develop predictive models for the abovementioned settings. Higher specificities were achieved with the SVM models. However, sensitivities for the SVM models were much lower than those for the ensembles models. The decision tree models demonstrated significantly inferior performance in terms of almost all the sensitivities, specificities, and AUCs. These results suggest that the ensemble models are able to yield excellent performance while maintaining a good balance between sensitivity and specificity in the prediction of CPP.

Feature Importance

We computed the feature importance score for all the 19 variables to identify important features used by the models. The

importance of each feature calculated by the models is plotted in Figure 3. In both models, the most important predictive variable was LH level, followed by IGF-I and FSH levels. The fourth most important feature for the random forest model was height, which ranked fifth in the XGBoost model. Prolactin is the fourth most important feature of XGBoost, but it contributed only a little to random forest. These data suggest that different machine learning algorithms attach importance to different combination lists of variables, although they yield similar predictive performance.

In order to further verify the importance of the top 3 features, we constructed XGBoost models with these features individually or in combination (Table 2). As expected, the models using one, two, or three features had poorer performances than the models using all features. The results showed that the higher the feature ranked, the better the corresponding model performed. Interestingly, LH alone or together with IGF-I and FSH is sufficient to predict a response to the GnRHa test with a fairly good performance and an AUC between 0.84 and 0.86. These data support the results from the feature importance calculations.

Figure 2. ROC curves for classifiers with 19 variables for 1757 patients and 25 variables for 436 patients. ROC: receiver operating curve; AUC: area under ROC.

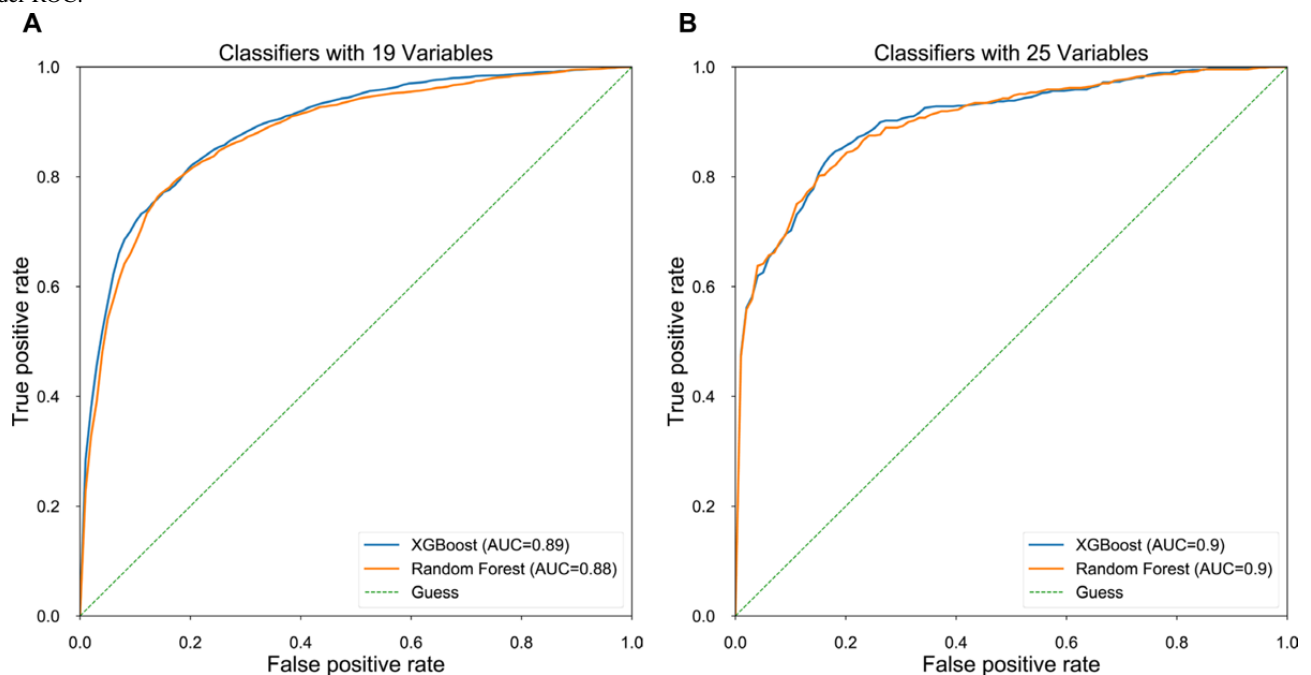
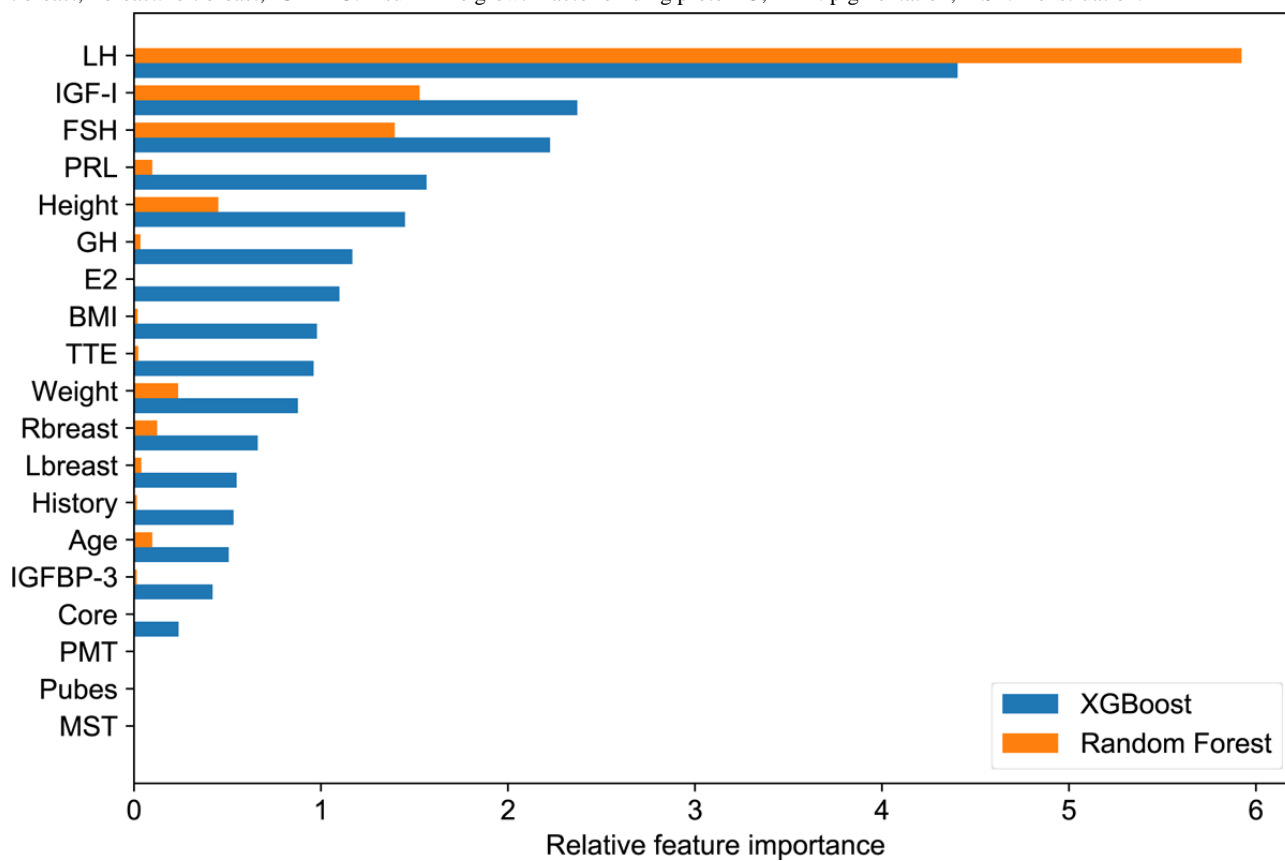


Figure 3. Feature importance ranking for 19 variables in two classifiers calculated by the models. LH: luteinizing hormone; IGF-I: insulin-like growth factor-I; FSH: follicle-stimulation hormone; PRL: prolactin; GH: growth hormone; E2: estradiol; BMI: body mass index; TTE: testosterone; Rbreast: right breast; Lbreast: left breast; IGFBP-3: insulin-like growth factor binding protein-3; PMT: pigmentation; MST: menstruation.

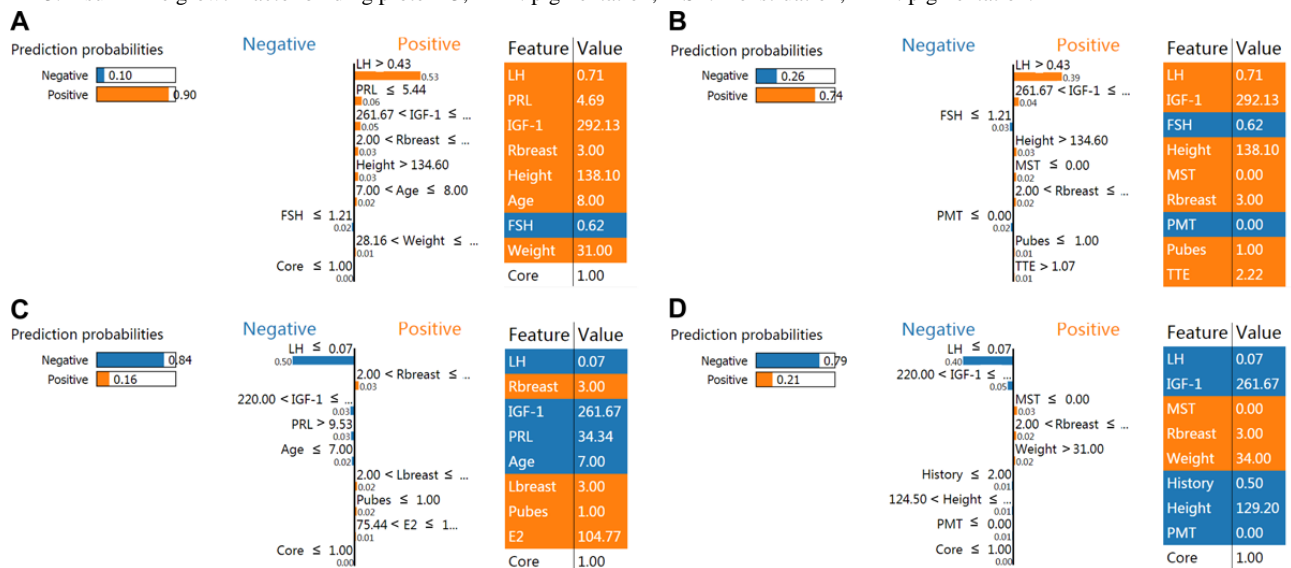


Local Interpretable Model-Agnostic Explanations for Interpretation

A representative set of 200 samples, which accounted for more than 50% of the test set and were enough to be the budget size of individual instances to understand a model, were selected with the submodular pick method [27] for the 19-variable models. LIME was then applied to investigate feature contributions for each prediction. Results with top 9 features are presented in Figure 4 for one positive sample and one negative sample (more representative samples can be seen in Multimedia Appendix 1). In Figure 4, XGBoost predicts an

instance where CPP positively responds to the GnRH α test with a probability of 90%. Only the feature growth hormone supports the negative prediction, whereas LH, prolactin, IGF-I, and information about body development support the positive prediction. This makes sense in the clinical diagnosis of CPP and reveals that we can trust our prediction models to a certain extent. In Figure 4, left and right breast at Tanner stage 3, FSH level > 2 IU/L, and several other features support the positive prediction with a probability of 16%. LH level of 0.07 IU/L, prolactin level > 9.53 ng/mL, IGF-I level > 220 ng/mL, and age < 7 years contribute to the negative prediction with a probability of 84%. Similar results are observed in Figure 4.

Figure 4. Results of LIME with XGBoost and Random Forest classifiers applied to one positive (A, B) and one negative (C, D) instance. The left sides are for XGBoost, and the right for Random Forest. Blue color is for the negative instance and orange is for the positive instance. The first column represents the prediction probabilities of negative and positive results achieved from classifiers. The second column shows the features' contributions to the probability. Only the top nine features are displayed for clarity. The third column displays the original data values. LIME: local interpretable model-agnostic explanations; XGBoost: extreme gradient boosting; LH: luteinizing hormone; IGF-I: insulin-like growth factor-I; FSH: follicle-stimulation hormone; PRL: prolactin; GH: growth hormone; E2: estradiol; BMI: body mass index; TTE: testosterone; Rbreast: right breast; Lbreast: left breast; IGFBP-3: insulin-like growth factor binding protein-3; PMT: pigmentation; MST: menstruation; PMT: pigmentation.



Discussion

Overview

CPP mimics pubertal development ahead of time at an inappropriate chronological age. It requires timely detection and treatment in case of physical and physiological effect on girls. The GnRHa-stimulation test is expensive and time consuming and causes discomfort to patients. Here, we applied machine learning algorithms to multiple clinical variables and built two tree-based ensemble learning classifiers for the prediction of response to the GnRHa-stimulation test. Both the XGBoost and random forest models achieved good performance in distinguishing between positive and negative responses, with the AUC ranging from 0.88 to 0.90, the sensitivity ranging from 77.91% to 77.94%, and the specificity ranging from 84.32% to 87.66%.

Comparisons with Previous Models

Several previous models focused on determining optimal blood sampling time points or appropriate cut-off values to simplify the stimulation test. Kandemir et al [29] found that a single sample of LH tested at the 40th minute after stimulation with a cut-off of 5 IU/L could yield 98% sensitivity and 100% specificity in the diagnosis of CPP. Yazdani et al [30] showed that an LH concentration > 5 IU/L at 3 hours has optimal sensitivity (83%) and specificity (97%). In the study of Çatlı et al [7], 100% sensitivity and 84% specificity were obtained using a cut-off value > 0.24 for peak LH/FSH ratio in girls. Although these models performed better than our models, they had to be used after stimulation and therefore could not avoid the disadvantages of the GnRH/GnRHa test completely.

Some models used only the basal sex hormone level. Yazdani et al [30] found that a basal LH level of >0.1 IU/L, a basal LH/FSH ratio >1, and basal estradiol level ≥1.5 ng/dL in girls

have low sensitivity (10%-67%) but excellent specificity (94%-100%). Çatlı et al [7] also reported models with the basal FSH or LH levels and achieved a sensitivity of 71% and a specificity of 68% or 64%. Pasternak et al [3] reported that basal LH levels ≤ 0.1 IU/L were sufficient to rule out positive response to the GnRH test with a specificity of 94% but a sensitivity of only 64% in girls. In another model [4], the basal LH level with a cut-off value of 0.35 IU/L was associated with a sensitivity of 63.96% and a specificity of 76.3% based on the ROC with an AUC of 0.77. These results varied a lot due to the different settings and sample sizes. In this study, our models showed better performance with more features before stimulation and a larger homogeneous population, which is the largest population in such a study to our knowledge.

Predictive Features

Based on our machine learning models, basal LH, IGF-I, and FSH levels are predictive factors with top ranks for the feature importance in both models. Previous studies have demonstrated that the measurement of LH could be better than that of other sex hormones for initial evaluation of suspected puberty [3,8]. In our study, the LH level ranked first and was much more important than other variables. Besides LH, another indicator monitored in the stimulation test, FSH, was also selected by the models as the third most-important variable. Obviously, LH and FSH are important to CPP because they are biomarkers of the hypothalamic-pituitary-gonadal axis activation, which is the essence of CPP. IGF-I, which is the second most important variable in our models, is reportedly involved in GnRH regulation [31,32] and is increasingly expressed in girls with CPP [9,5]. Animal studies showed that the IGF-I signaling pathways play important roles in the timing of puberty in girls [32]. Although IGF-I has not been considered in previous models, our study suggests that IGF-I may be a valuable marker for diagnosing CPP.

Several studies [12,33,34] suggested that image reports like pelvic ultrasound and radiography of the hand have adjunct diagnostic values in CPP diagnosis but provide no reliable differentiation alone. Here, we found that adding features from the image reports improved the prediction results. Performance of models built based on 1757 samples was better than that based on 436 samples, suggesting a sample size effect. Interestingly, in the case of 436 samples with additional six image variables, the aforementioned sample size effect was balanced. This suggests that more samples with image features will produce better results. Thus, medical image examinations like bone age radiography should be considered before the GnRHa-stimulation test for girls with suspected CPP.

Interpretations of Models

We noticed that more variables were assigned with a moderate value of feature importance in the XGBoost model than those in the random forest model. This is reasonable when considering the different algorithms the two models used for prediction and importance evaluation. In XGBoost, trees are sequentially built in a boosting manner to enhance the overall performance. The estimates of feature importance are provided explicitly with the frequency that the feature is selected as a tree node from a trained predictive model. In contrast, trees are trained parallelly in a bootstrapping way in random forest to vote for the final decision. The feature importance is estimated implicitly through permuting the feature's values and calculating the change of the model's prediction error. Obviously, XGBoost includes each contribution of each feature to each tree into the feature importance, whereas random forest only evaluates each feature globally without specific contributions. It should be noted that different combinations of variables may produce models with similar predictive accuracy, relating to the uncertainty analysis of the solutions in any decision-making problem [35-37]. This is not rare in machine learning models in medicine [38-40]. Moreover, the most important features in the clinic such as basal LH, IGF-I, and FSH levels were all sorted out by both models, demonstrating that they are both reliable and effective in predicting response to the GnRHa test.

In order to provide endocrinology physicians a trustworthy insight into the prediction models, we also used LIME to show each feature's contribution to predicting probabilities reasonably. The most important features used by the models for individual prediction have been proven to be significant in the clinic [3-5,7,9,10], demonstrating that our models are credible. This will greatly increase the interpretability of the machine learning models and make it convenient for individualized diagnosis in the clinic.

Limitations

There are some limitations to this study. First, growth velocity is specially related to physical development. Due to the lack of height growth rate and weight growth rate, we did not include growth velocity in our feature set. For further research, we will focus more on medical imaging and growth velocity to identify their diagnostic value with CPP. Second, our work included only girls with suspected CPP from a single center in China. The prediction models in this study may not be suitable for the population in other districts or countries. Third, manual inspection of values extracted through regular expression matching from free text could reduce errors to improve the model performance. However, this adds a considerable amount of work and thus reduces the scalability of the model. We are improving the matching algorithm with the manually inspected data to increase the level of model automation. Finally, features generated from laboratory results are more complete than those extracted from free text, which may affect the rank of feature importance. More efforts are required to enhance the data quality of unstructured features in the future.

Conclusions

Our study is the first one to apply both machine learning algorithms and the explanation method to the diagnosis of CPP. Our models can predict the response to the stimulation test before injection of GnRHa in girls who are suspected of having CPP and thus may be used as a prescreening tool to help physicians make decisions in conjunction with the GnRHa-stimulation test.

Acknowledgments

This work was supported by Guangzhou Institute of Pediatrics, Guangzhou Women and Children's Medical Center (Grant #KCP-2016-002).

Authors' Contributions

HL and XL proposed the project and provided supervision. LP is responsible for data collection, project implementation, and manuscript drafting and revision. GL implemented the project and performed manuscript editing. XM provided clinical guidance and revised the manuscript. HL conducted part of the data collection and statistical analysis. JZ helped collect part of the data. All authors reviewed the manuscript in its final form.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results of LIME with XGBoost and Random Forest classifiers applied to four positive and four negative instances. LIME: local interpretable model-agnostic explanations.

[PDF File (Adobe PDF File), 267KB - [medinform_v7i1e11728_app1.pdf](#)]

References

1. Abreu AP, Kaiser UB. Pubertal development and regulation. *Lancet Diabetes Endocrinol* 2016 Mar;4(3):254-264 [FREE Full text] [doi: [10.1016/S2213-8587\(15\)00418-0](#)] [Medline: [26852256](#)]
2. Houk C, Kunselman AR, Lee PA. The diagnostic value of a brief GnRH analogue stimulation test in girls with central precocious puberty: a single 30-minute post-stimulation LH sample is adequate. *J Pediatr Endocrinol Metab* 2008 Dec;21(12):1113-1118. [Medline: [19189683](#)]
3. Pasternak Y, Friger M, Loewenthal N, Haim A, Hershkovitz E. The utility of basal serum LH in prediction of central precocious puberty in girls. *Eur J Endocrinol* 2012 Feb;166(2):295-299 [FREE Full text] [doi: [10.1530/EJE-11-0720](#)] [Medline: [22084156](#)]
4. Ding Y, Li J, Yu Y, Yang P, Li H, Shen Y, et al. Evaluation of basal sex hormone levels for activation of the hypothalamic-pituitary-gonadal axis. *J Pediatr Endocrinol Metab* 2018 Mar 28;31(3):323-329. [doi: [10.1515/jpem-2017-0124](#)] [Medline: [29466239](#)]
5. Sørensen K, Aksglaede L, Petersen JH, Andersson AM, Juul A. Serum IGF1 and insulin levels in girls with normal and precocious puberty. *Eur J Endocrinol* 2012 May;166(5):903-910. [doi: [10.1530/EJE-12-0106](#)] [Medline: [22379117](#)]
6. Houk C, Kunselman AR, Lee PA. Adequacy of a single unstimulated luteinizing hormone level to diagnose central precocious puberty in girls. *Pediatrics* 2009 Jun;123(6):e1059-e1063. [doi: [10.1542/peds.2008-1180](#)] [Medline: [19482738](#)]
7. Çatlı G, Erdem P, Anık A, Abacı A, Böber E. Clinical and laboratory findings in the differential diagnosis of central precocious puberty and premature thelarche. *Turk Pediatri Ars* 2015 Mar;50(1):20-26 [FREE Full text] [doi: [10.5152/tpa.2015.2281](#)] [Medline: [26078693](#)]
8. Mogensen SS, Aksglaede L, Mouritsen A, Sørensen K, Main KM, Gideon P, et al. Diagnostic work-up of 449 consecutive girls who were referred to be evaluated for precocious puberty. *J Clin Endocrinol Metab* 2011 May;96(5):1393-1401. [doi: [10.1210/jc.2010-2745](#)] [Medline: [21346077](#)]
9. Nam H, Rhie YJ, Son CS, Park SH, Lee K. Factors to predict positive results of gonadotropin releasing hormone stimulation test in girls with suspected precocious puberty. *J Korean Med Sci* 2012 Feb;27(2):194-199 [FREE Full text] [doi: [10.3346/jkms.2012.27.2.194](#)] [Medline: [22323868](#)]
10. Suh J, Choi MH, Kwon AR, Kim YJ, Jeong JW, Ahn JM, et al. Factors that predict a positive response on gonadotropin-releasing hormone stimulation test for diagnosing central precocious puberty in girls. *Ann Pediatr Endocrinol Metab* 2013 Dec;18(4):202-207 [FREE Full text] [doi: [10.6065/apem.2013.18.4.202](#)] [Medline: [24904878](#)]
11. Xu Y, Li GM, Li Y. Advanced bone age as an indicator facilitates the diagnosis of precocious puberty. *J Pediatr (Rio J)* 2018;94(1):69-75 [FREE Full text] [doi: [10.1016/j.jpmed.2017.03.010](#)] [Medline: [28866322](#)]
12. Lee SH, Joo EY, Lee J, Jun Y, Kim M. The Diagnostic Value of Pelvic Ultrasound in Girls with Central Precocious Puberty. *Chonnam Med J* 2016 Jan;52(1):70-74 [FREE Full text] [doi: [10.4068/cmj.2016.52.1.70](#)] [Medline: [26866003](#)]
13. Lee H, Choi S, Fan D, Jang K, Kim M, Hwang C. Evaluation of characteristics of the craniofacial complex and dental maturity in girls with central precocious puberty. *Angle Orthod* 2018 Apr 30. [doi: [10.2319/112317-809.1](#)] [Medline: [29708396](#)]
14. Baik JS, Choi JW, Kim SJ, Kim JH, Kim S, Kim JH. Predictive Value of Dental Maturity for a Positive Gonadotropin-Releasing Hormone Stimulation Test Result in Girls with Precocious Puberty. *J Korean Med Sci* 2017 Feb;32(2):296-302 [FREE Full text] [doi: [10.3346/jkms.2017.32.2.296](#)] [Medline: [28049241](#)]
15. Subspecialty Group of Endocrinologic, Hereditary and Metabolic Diseases, the Society of Pediatrics, Chinese Medical Association, Editorial Board, Chinese Journal of Pediatrics. [Consensus statement For the diagnosis and treatment of central precocious puberty (2015)]. *Zhonghua Er Ke Za Zhi* 2015 Jun;53(6):412-418. [Medline: [26310550](#)]
16. Carel J, Eugster EA, Rogol A, Ghizzoni L, Palmert MR, ESPE-LWPES GnRH Analogs Consensus Conference Group, et al. Consensus statement on the use of gonadotropin-releasing hormone analogs in children. *Pediatrics* 2009 Apr;123(4):e752-e762. [doi: [10.1542/peds.2008-1783](#)] [Medline: [19332438](#)]
17. Ibáñez L, Potau N, Zampolli M, Virdis R, Gussinyé M, Carrascosa A, et al. Use of leuprolide acetate response patterns in the early diagnosis of pubertal disorders: comparison with the gonadotropin-releasing hormone test. *J Clin Endocrinol Metab* 1994 Jan;78(1):30-35. [doi: [10.1210/jcem.78.1.7507123](#)] [Medline: [7507123](#)]
18. Bayley N, Pinneau SR. Tables for predicting adult height from skeletal age: Revised for use with the greulich-pyle hand standards. *The Journal of Pediatrics* 1952 Apr;40(4):423-441. [doi: [10.1016/S0022-3476\(52\)80205-7](#)]
19. Python. URL: <https://www.python.org/> [accessed 2019-01-22] [WebCite Cache ID 75cTVrNhZ]
20. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016 Aug 13 Presented at: KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA, USA. [doi: [10.1145/2939672.2939785](#)]
21. Torlay L, Perrone-Bertolotti M, Thomas E, Baciú M. Machine learning-XGBoost analysis of language networks to classify patients with epilepsy. *Brain Inform* 2017 Sep;4(3):159-169 [FREE Full text] [doi: [10.1007/s40708-017-0065-7](#)] [Medline: [28434153](#)]

22. Yao D, Calhoun VD, Fu Z, Du Y, Sui J. An ensemble learning system for a 4-way classification of Alzheimer's disease and mild cognitive impairment. *J Neurosci Methods* 2018 May 15;302:75-81. [doi: [10.1016/j.jneumeth.2018.03.008](https://doi.org/10.1016/j.jneumeth.2018.03.008)] [Medline: [29578038](https://pubmed.ncbi.nlm.nih.gov/29578038/)]
23. Liaw A, Wiener M. R News. 2002 Dec. Classification and regression by randomForest URL: https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf [accessed 2019-01-22] [WebCite Cache ID 75cYZwTh7]
24. Hearst M, Dumais S, Osuna E, Platt J, Scholkopf B. Support vector machines. *IEEE Intell Syst Their Appl* 1998 Jul 10;13(4):18-28. [doi: [10.1109/5254.708428](https://doi.org/10.1109/5254.708428)] [Medline: [21889629](https://pubmed.ncbi.nlm.nih.gov/21889629/)]
25. Safavian S, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern* 1991;21(3):660-674. [doi: [10.1109/21.97458](https://doi.org/10.1109/21.97458)]
26. Xu Q, Liang Y. Monte Carlo cross validation. *Chemometrics and Intelligent Laboratory Systems* 2001 Apr;56(1):1-11. [doi: [10.1016/S0169-7439\(00\)00122-2](https://doi.org/10.1016/S0169-7439(00)00122-2)]
27. Ribeiro MT, Singh S, Guestrin C. Why should i trust you? Explaining the predictions of any classifier. 2016 Aug Presented at: KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, California, USA p. E. [doi: [10.1145/2939672.2939778](https://doi.org/10.1145/2939672.2939778)]
28. Liu N, Kumara S, Reich E. Explainable data-driven modeling of patient satisfaction survey data. : IEEE; 2017 Dec Presented at: 2017 IEEE International Conference on Big Data; 11-14 Dec. 2017; Boston, MA, USA. [doi: [10.1109/BigData.2017.8258391](https://doi.org/10.1109/BigData.2017.8258391)]
29. Kandemir N, Demirbilek H, Özön ZA, Gönc N, Alikeşifoğlu A. GnRH stimulation test in precocious puberty: single sample is adequate for diagnosis and dose adjustment. *J Clin Res Pediatr Endocrinol* 2011 Mar;3(1):12-17 [FREE Full text] [doi: [10.4274/jcrpe.v3i1.03](https://doi.org/10.4274/jcrpe.v3i1.03)] [Medline: [21448328](https://pubmed.ncbi.nlm.nih.gov/21448328/)]
30. Yazdani P, Lin Y, Raman V, Haymond M. A single sample GnRHa stimulation test in the diagnosis of precocious puberty. *Int J Pediatr Endocrinol* 2012 Jul 18;2012(1):23 [FREE Full text] [doi: [10.1186/1687-9856-2012-23](https://doi.org/10.1186/1687-9856-2012-23)] [Medline: [22809285](https://pubmed.ncbi.nlm.nih.gov/22809285/)]
31. Wolfe A, Divall S, Wu S. The regulation of reproductive neuroendocrine function by insulin and insulin-like growth factor-1 (IGF-1). *Front Neuroendocrinol* 2014 Oct;35(4):558-572 [FREE Full text] [doi: [10.1016/j.yfrne.2014.05.007](https://doi.org/10.1016/j.yfrne.2014.05.007)] [Medline: [24929098](https://pubmed.ncbi.nlm.nih.gov/24929098/)]
32. Divall SA, Williams TR, Carver SE, Koch L, Brüning JC, Kahn CR, et al. Divergent roles of growth factors in the GnRH regulation of puberty in mice. *J Clin Invest* 2010 Aug;120(8):2900-2909 [FREE Full text] [doi: [10.1172/JCI41069](https://doi.org/10.1172/JCI41069)] [Medline: [20628204](https://pubmed.ncbi.nlm.nih.gov/20628204/)]
33. de Vries L, Phillip M. Role of pelvic ultrasound in girls with precocious puberty. *Horm Res Paediatr* 2011 Feb;75(2):148-152. [doi: [10.1159/000323361](https://doi.org/10.1159/000323361)] [Medline: [21228561](https://pubmed.ncbi.nlm.nih.gov/21228561/)]
34. Eksioğlu A, Yılmaz S, Cetinkaya S, Cinar G, Yıldız YT, Aycan Z. Value of pelvic sonography in the diagnosis of various forms of precocious puberty in girls. *J Clin Ultrasound* 2013 Feb;41(2):84-93. [doi: [10.1002/jcu.22004](https://doi.org/10.1002/jcu.22004)] [Medline: [23124596](https://pubmed.ncbi.nlm.nih.gov/23124596/)]
35. Fernández Martínez J, Fernández Muñoz M, Tompkins M. On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics* 2012 Jan;77(1):W1-W15. [doi: [10.1190/geo2011-0341.1](https://doi.org/10.1190/geo2011-0341.1)]
36. Polikar R. Ensemble based systems in decision making. *IEEE Circuits Syst Mag* 2006 Sep;6(3):21-45. [doi: [10.1109/MCAS.2006.1688199](https://doi.org/10.1109/MCAS.2006.1688199)]
37. Fernández-Martínez J, Fernández-Muñoz Z, Palleró J, Pedruelo-González L. From Bayes to Tarantola: New insights to understand uncertainty in inverse problems. *Journal of Applied Geophysics* 2013 Nov;98:62-72. [doi: [10.1016/j.jappgeo.2013.07.005](https://doi.org/10.1016/j.jappgeo.2013.07.005)]
38. deAndrés-Galiana EJ, Fernández-Martínez JL, Luaces O, Del CJJ, Huergo-Zapico L, Acebes-Huerta A, et al. Analysis of clinical prognostic variables for Chronic Lymphocytic Leukemia decision-making problems. *J Biomed Inform* 2016 Apr;60:342-351 [FREE Full text] [doi: [10.1016/j.jbi.2016.02.017](https://doi.org/10.1016/j.jbi.2016.02.017)] [Medline: [26956213](https://pubmed.ncbi.nlm.nih.gov/26956213/)]
39. Kosmicki J, Sochat V, Duda M, Wall DP. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. *Transl Psychiatry* 2015 Feb 24;5:e514 [FREE Full text] [doi: [10.1038/tp.2015.7](https://doi.org/10.1038/tp.2015.7)] [Medline: [25710120](https://pubmed.ncbi.nlm.nih.gov/25710120/)]
40. Liu G, Xu Y, Wang X, Zhuang X, Liang H, Xi Y, et al. Developing a Machine Learning System for Identification of Severe Hand, Foot, and Mouth Disease from Electronic Medical Record Data. *Sci Rep* 2017 Nov 27;7(1):16341 [FREE Full text] [doi: [10.1038/s41598-017-16521-z](https://doi.org/10.1038/s41598-017-16521-z)] [Medline: [29180702](https://pubmed.ncbi.nlm.nih.gov/29180702/)]

Abbreviations

- AUC:** area under receiver operating characteristic
- BMI:** body mass index
- CPP:** central precocious puberty
- FSH:** follicle-stimulation hormone
- GnRH:** gonadotropin releasing hormone
- GnRHa:** gonadotropin releasing hormone analogues
- IGF-I:** insulin-like growth factor-I
- IGFBP-3:** insulin-like growth factor binding protein-3

LH: luteinizing hormone
LIME: local interpretable model-agnostic explanations
MCCV: Monte Carlo cross-validation
OOB: out-of-bag
ROC: receiver operating characteristic
SVM: supported vector machines
XGBoost: extreme gradient boosting

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by P Cai, J Chen; comments to author 06.09.18; revised version received 10.10.18; accepted 09.12.18; published 12.02.19.

Please cite as:

Pan L, Liu G, Mao X, Li H, Zhang J, Liang H, Li X

Development of Prediction Models Using Machine Learning Algorithms for Girls with Suspected Central Precocious Puberty: Retrospective Study

JMIR Med Inform 2019;7(1):e11728

URL: <http://medinform.jmir.org/2019/1/e11728/>

doi: [10.2196/11728](https://doi.org/10.2196/11728)

PMID: [30747712](https://pubmed.ncbi.nlm.nih.gov/30747712/)

©Liyan Pan, Guangjian Liu, Xiaojian Mao, Huixian Li, Jiexin Zhang, Huiying Liang, Xiuzhen Li. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 12.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Medication Use for Childhood Pneumonia at a Children's Hospital in Shanghai, China: Analysis of Pattern Mining Algorithms

Chunlei Tang^{1*}, PhD; Huajun Sun^{2*}, PhD; Yun Xiong^{3*}, PhD; Jiahong Yang⁴, MPH; Christopher Vitale⁵, PharmD; Lu Ruan³, MSc; Angela Ai¹, BA; Guangjun Yu², PhD; Jing Ma⁶, MD, PhD; David Bates¹, MD, MSc

¹Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

²Children's Hospital of Shanghai, Shanghai Jiaotong University School of Medicine, Shanghai, China

³Shanghai Key Laboratory of Data Science, School of Computer Science, Fudan University, Shanghai, China

⁴Shanghai Shenkang Hospital Development Center, Shanghai, China

⁵Clinical Informatics for the Integrated Health Model Initiative, American Medical Association, Chicago, IA, United States

⁶Department of Population Medicine, Harvard Medical School, Boston, MA, United States

*these authors contributed equally

Corresponding Author:

Guangjun Yu, PhD

Children's Hospital of Shanghai

Shanghai Jiaotong University School of Medicine

No 24, Lane 1400 West Beijing Road,

Shanghai, 200040

China

Phone: 86 18917762998

Email: gju@shchildren.com.cn

Abstract

Background: Pattern mining utilizes multiple algorithms to explore objective and sometimes unexpected patterns in real-world data. This technique could be applied to electronic medical record data mining; however, it first requires a careful clinical assessment and validation.

Objective: The aim of this study was to examine the use of pattern mining techniques on a large clinical dataset to detect treatment and medication use patterns for childhood pneumonia.

Methods: We applied 3 pattern mining algorithms to 680,138 medication administration records from 30,512 childhood inpatients with diagnosis of pneumonia during a 6-year period at a children's hospital in China. Patients' ages ranged from 0 to 17 years, where 37.53% (11,453/30,512) were 0 to 3 months old, 86.55% (26,408/30,512) were under 5 years, 60.37% (18,419/30,512) were male, and 60.10% (18,338/30,512) had a hospital stay of 9 to 15 days. We used the FP-Growth, PrefixSpan, and USpan pattern mining algorithms. The first 2 are more traditional methods of pattern mining and mine a complete set of frequent medication use patterns. PrefixSpan also incorporates an administration sequence. The newer USpan method considers *medication utility*, defined by the dose, frequency, and timing of use of the 652 individual medications in the dataset. Together, these 3 methods identified the top 10 patterns from 6 age groups, forming a total of 180 distinct medication combinations. These medications encompassed the top 40 (73.66%, 500,982/680,138) most frequently used medications. These patterns were then evaluated by subject matter experts to summarize 5 medication use and 2 treatment patterns.

Results: We identified 5 medication use patterns: (1) antiasthmatics and expectorants and corticosteroids, (2) antibiotics and (antiasthmatics or expectorants or corticosteroids), (3) third-generation cephalosporin antibiotics with (or followed by) traditional antibiotics, (4) antibiotics and (medications for enteritis or skin diseases), and (5) (antiasthmatics or expectorants or corticosteroids) and (medications for enteritis or skin diseases). We also identified 2 frequent treatment patterns: (1) 42.89% (291,701/680,138) of specific medication administration records were of intravenous therapy with antibiotics, diluents, and nutritional supplements and (2) 11.53% (78,390/680,138) were of various combinations of inhalation of antiasthmatics, expectorants, or corticosteroids. Fleiss kappa for the subject experts' evaluation was 0.693, indicating moderate agreement.

Conclusions: Utilizing a pattern mining approach, we summarized 5 medication use patterns and 2 treatment patterns. These warrant further investigation.

KEYWORDS

drug therapy; combination; computer-assisted; pattern recognition; data mining; precision medicine; childhood pneumonia; hospital

Introduction

Childhood pneumonia remains the single largest cause of death in young children worldwide [1,2]. According to a recent World Health Organization (WHO) report, an estimated 922,000 children under the age of 5 passed away because of pneumonia in 2015 alone, accounting for 16% of all deaths in this age group [2,3]. In China, it is an especially grave concern, driven in part by environmental pollution [3] and abuse of antibiotics [4]. For virtually all high-mortality settings, the WHO's case management plan uses algorithms as the basis for pneumonia management [5]. Globally, there are many clinical recommendations for treating pediatric pneumonia, but some guidelines are outdated and can be vague [6]. Variation in treatment regimens within the same clinic could indicate poor clinical practice leading to increased recurrence rates that are more likely to have complications [7,8]. Multiple medication therapies are common practice and may provide more effective treatment with a lower concentration of individual components lessening the risk of side effects and toxicity [9,10]. What is missing, however, is the knowledge derived directly from real-world clinical practice.

Electronic medical record (EMR) data represent a rich source of information that includes 2 types of clinical knowledge [8,11,12]. First, it reflects the medical knowledge— and specialty-based clinical practice of a group of doctors within a certain time period. This encompasses medical solutions originating from various clinical standards developed by physicians as a professional group. Second, it represents the clinical experience of an individual practicing clinician and their prescribing style, as well as external factors such as pricing and regulation. Meaningful use of these data can yield information that can guide clinical practice. Challenges of using EMR data, however, have included the large data volume and the substantial variations between different EMR systems.

Data mining techniques have great potential for use in exploring EMR data. Pattern mining is an important subfield of data mining. Its goal is to find all possible salient and persistent patterns in a dataset [13,14], including, but not limited to, direct or indirect associations, trends, periodic patterns, sequential rules, and high-utility patterns of real-life events. Pattern mining is an overloaded term that is used to refer to different technologies in different domains. These techniques have been used in economics, for example, to analyze the stock market and supermarket sales, but have rarely been used in medical research and, specifically, aiming to a variable length of hospital stay is the first time. In one context, it means statistical properties for handling continuous attributes, and in another context, it refers to an ordinal relation, usually based on temporal or spatial precedence that exists among events occurring in the data. Furthermore, the tools and evaluation methods used in each context are different. We aimed to consider the sequential

information that may be valuable for identifying recurring features of a dynamic system or predicting future occurrences of certain events. In addition, most of the previous research on discovering medication-using patterns involves text-based methods and is limited to one or two medications [15,16]. In this study, we explored the utility of a pattern mining approach to the entire record of inpatient EMR-based medications administered for childhood pneumonia in a large children's hospital in China across several years.

Methods

Study Population

The Children's Hospital of Shanghai (CHS), one of the top comprehensive children's hospitals in China, admits approximately 5000 inpatients annually from different parts of the country and internationally. Formerly known as Underprivileged Children's Hospital, it is one of the oldest children's hospitals in Asia.

We extracted 680,138 inpatient medication administration records from 30,512 childhood patients with an initial diagnosis of pneumonia, from January 1, 2010, to December 31, 2015, at the CHS, China. The EMR dataset contains 18,419 males and 12,093 females, among which 60.10% (18,338/30,512) had a hospital stay of 9 to 15 days. If rehospitalized, the patients had 1 record for each hospitalization. Raw data were deidentified. We pulled age and sex, initial diagnosis, and admission and discharge dates, as well as routes and doses of administered medications. According to Chinese legislation, ethical approval from the regional ethical review board is not needed for this type of study of deidentified EMR data. The data used in this study were anonymous. Although ethical approval from the regional ethical review board is not needed for this type of study of deidentified EMRs according to Chinese legislation, we still applied for and received approval from the Institutional Review Board of the CHS.

Data Preparation

We first merged similar routes of administration for the same medication, and then converted commonly used medications to their shortened reference names as shown in [Multimedia Appendix 1](#). We then removed diluents (or carriers) from the dataset after our expert panel determined that these were not intended to have a therapeutic effect. The final dataset contains 652 individual medications (including traditional Chinese patent medicine [17]), among which the top 40 medications appeared in 73.66% (500,982/680,138) of all the medication administration records ([Multimedia Appendix 1](#)).

Selected Pattern Mining Algorithms: FP-Growth, PrefixSpan, and USpan

Our initial goal was to generate sets of frequently appearing items (ie, frequent item sets) to better understand *medication*

use patterns. However, doing so would involve excluding too many temporal data points. An example of this temporal information is the time at which an item (eg, medication) was prescribed by a physician. Thus, while the classic frequency-based framework often leads to many patterns being identified, most of them are not informative enough for further clinical investigation. Recent efforts have been made to incorporate utility into the pattern selection framework, so that high-utility patterns, regardless of frequency, are mined. To obtain more reliable and relevant results, we first developed a unified platform to draw a parallel comparison among all medication use patterns produced by 3 pattern mining algorithms (FP-Growth, PrefixSpan, and USpan). We then asked the panel of subject matter experts to review these results based on their clinical expertise and to make recommendations for meaningful grouping. We also compared these results with what experts found using summary statistics in our previous work [18].

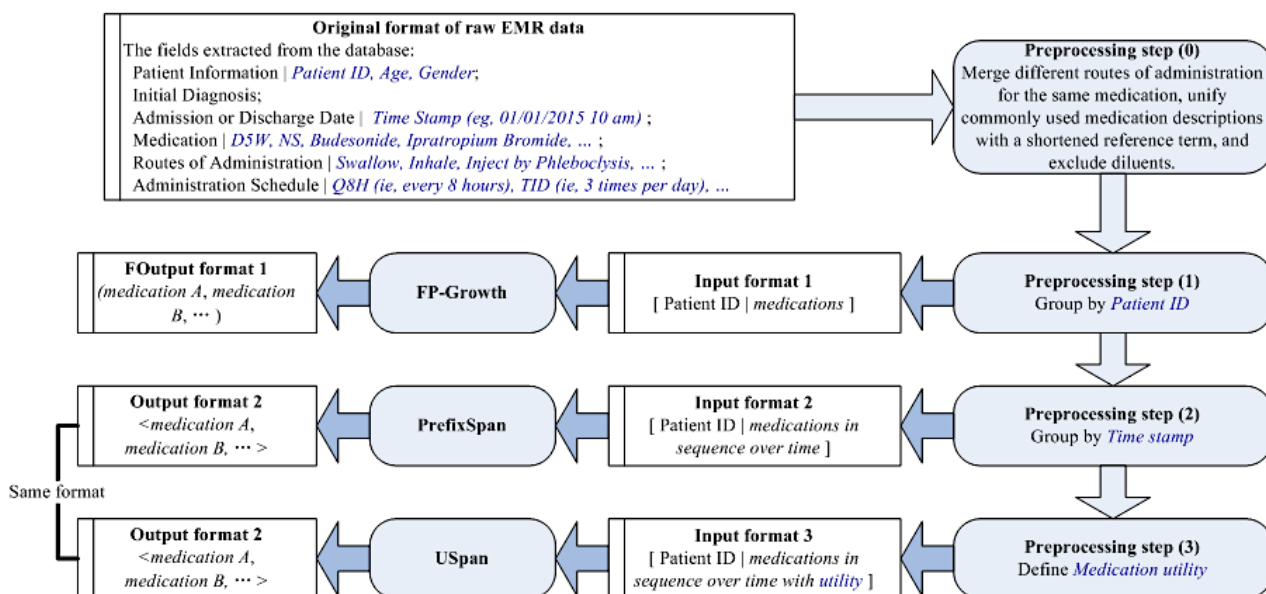
The advantage of progressively using 3 algorithms instead of applying only 1 is twofold: (1) the 3 methods allow for more comprehensive data mining that includes frequency, timing, and utility of dosing of medication administration and (2) it can reduce approach-specific limitations. FP-Growth and PrefixSpan are 2 classic pattern mining algorithms, which were commonly used in business. The earliest and most classic example of their use was examining the cosale of diapers and beer [19]. FP-Growth allows mining of a complete set of frequent patterns by pattern fragment growth without using candidate generations [20]. PrefixSpan offers ordered growth and reduced projected databases over a particular interval [21]. However, the 2 algorithms can only find patterns that appear with high frequency. For example, if there is a medication use pattern of

(a, b) where one of the medications (ie, a or b) was administered with low frequency, the 2 algorithms would not be able to discover this combination. It is possible that some interesting patterns (eg, a medication prescribed at a low frequency but at a high dose) may be filtered out of the results because of the medication’s low frequency of appearance. To address this problem, we used the more advanced USpan algorithm (first proposed by Yin et al [22]) and proposed a new definition called *medication utility*. The term *utility* originated from economics and considers quantities, profits, and time orders of items simultaneously [23]. The method is unique in its business consideration in dollar value for customers in financial markets. We define *medication utility* using both the dosage and frequency of medication administration over time to obtain more detailed evidence of medication use. USpan reflects a recent effort to incorporate utility into the pattern selection framework, so that high-utility (frequent or infrequent) patterns are mined which address some of the concerns involved in exploratory factor analysis, such as dollar value associated with each pattern. To achieve this, we implemented the USpan algorithm (see [Multimedia Appendix 2](#) with its explanation).

Application of Pattern Mining Algorithms to the Electronic Medical Record Data

We summarized our approach in [Figure 1](#), including (1) the original format of the raw EMR data, (2) a diagram outlining the 3-step data processing method, and (3) the steps taken by a particular algorithm, and its input and output formats. The thresholds used in the algorithms are as follows: the minimum support of both FP-Growth and PrefixSpan is 0.15, and the minimum utility of USpan is 30,000. We only utilized the top 10 results from our algorithmic outputs.

Figure 1. Application of algorithms to the dataset (medication A is a placeholder for a drug name and a indicates the frequency of appearance of medication A). EMR: electronic medical record; D5W: Dextrose 5% in Water; NS: normal saline; ID: identification.



Step 1: Group by Patient Identification

As FP-Growth disregards the quantity of items and time information, we first simplified the raw EMR data to a list of all the administered medications and their corresponding patients. FP-Growth outputs an unordered pattern represented by several medications enclosed in parentheses, for example, (*medication A*, *medication B*, and *medication C*). For example, consider 2 patients, one who used *medication A* on January 1 and *medication B* on January 2, and another who used *medication B* on January 10 and *medication A* on January 21. In this scenario, the output of FP-Growth would be (*medication A* and *medication B*), meaning that *medication A* and *medication B* were administered during the same hospitalization.

Step 2: Group by Time Stamp

To make a suitable input for PrefixSpan that includes temporality, we took the results from Step 1 and determined the sequence of medication administration over time. PrefixSpan outputs a temporal order enclosed in a pair of angled brackets. For example, the combination $\langle (\textit{medication A}, \textit{medication B}), \textit{medication C}, \dots \rangle$ means that there exists a temporal relation between (*medication A*, *medication B*) and *medication C*. Put another way, *medication A* and *medication B* were administered at the same time, and both were administered before *medication C*. For example, if a patient uses *medication A* for the first day, *medication B* for the first day, *medication D* for the fifth day, *medication A* for the eighth day, and *medication B* for the ninth day, while FP-Growth would output the grouping (*medication A*, *medication B*, *medication D*) without considering the repeat medications, PrefixSpan's output would be of the form $\langle (\textit{medication A}, \textit{medication B}), \textit{medication D}, \textit{medication A}, \textit{medication B} \rangle$.

Step 3: Define Medication Utility

We defined *medication utility* as follows. Let $I = \{i_1, i_2, \dots, i_n\}$ be the universal set of distinct medications. Each item $i_k \in I$ ($1 \leq k \leq n$) is associated with a utility value, denoted as $P(i_k)$, which shows the utility of dosing of each specific medication to treat a given disease. A medication usage is represented as an ordered pair (i_k, q_k) , where $i_k \in I$ is a medication and q_k is a positive number representing the quantity of I , which shows how much of this medication is taken (ie, dose). We thus define *medication utility* as $U(i_k, q_k) = P(i_k) \times q_k$ representing a single medication's dosage during one patient's hospitalization.

Since USpan takes both the time sequence and medication utility of each item into account, every administration record that occurred at the same time will be enclosed in a pair of parentheses, such as (*medication A*, a), in which *medication A* is a single drug and a is the frequency of *medication A*'s appearance. Each medication is attached to its utility, and each medication-utility pair is separated by time, indicated by a comma. USpan results are in the same output format as PrefixSpan.

Experts' Reviewing of the Results

To determine which machine-identified patterns most accurately reflect real-world clinical practice [24], we invited a panel of

subject matter experts including 3 physicians (MD), 3 pharmacists, and 2 researchers (1 DS and 1 MPH). The panelists were invited to review, adjusting process inputs and outputs if needed, and offer opinions on the clinical validity of the patterns obtained from the 3 algorithms. We then conducted a thorough literature review (by putting the drug names seen in a specific pattern into Google's search engine to find related papers) to further analyze the combinations that our panel considered clinically interesting. Our experts then conducted a final review of the results of the literature search to further validate the medication use patterns. We calculated Fleiss kappa to measure interrater reliability for the experts shown in [Multimedia Appendix 3](#).

Results

Study Population

The majority (86.55%, 26,408/30,512) of the study population ([Table 1](#)) was under 5 years of age. Among the under-5 population, 66.55% (20,305/30,512) were under 2 years of age and 37.53% (11,453/30,512) were newborns. There were more male (60.37%, 18,419/30,512) than female patients.

Simple Exploration of Frequency of Use of the Raw Electronic Medical Record Data

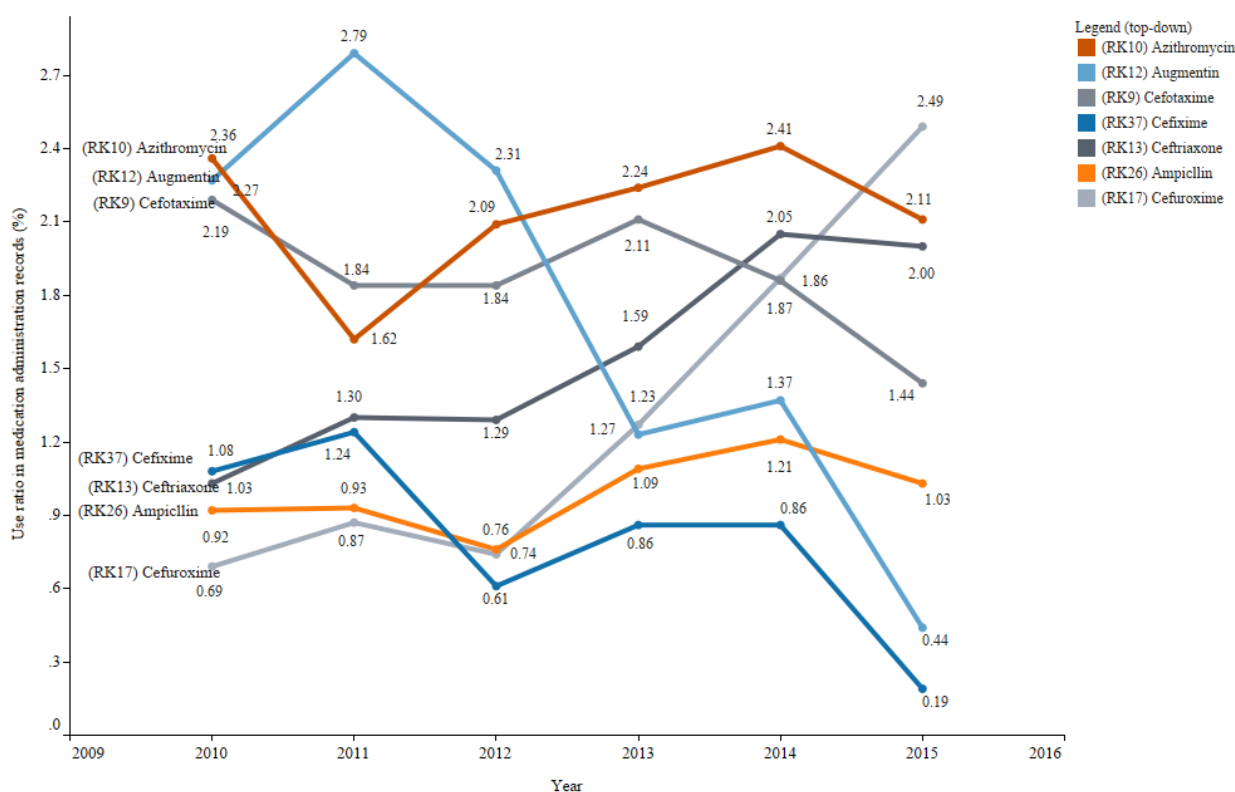
To better understand the raw EMR data, we first examined the top 40 most frequently prescribed medications across age groups and found that 2 diluents (D5W, Dextrose 5% in Water and NS, Normal Saline) were ranked the first- and second-most frequent ([Multimedia Appendix 4](#)). After removing all diluents from the dataset, some nutritional supplements, such as fat emulsion and vitamins, were ranked within top 10 medications for each of the 6 age groups ([Multimedia Appendix 5](#)). Besides diluents and nutritional supplements, antibiotics were also among the top 10 most frequently prescribed in each age group. These medications were present in 42.89% (291,701/680,138) of all specific medication administration records and they were all administered through intravenous (IV) therapy. Thus, *treatment pattern 1* is the combination of antibiotics, diluents, and nutritional supplements via IV therapy.

Due to the high frequency of antibiotic use, we then used statistics on individual medications to examine the frequency of use of the 7 commonly administered antibiotic monotherapies over the study period (2010 to 2015) and found a dramatically increased use of ceftriaxone (a third-generation cephalosporin antibiotic) and cefuroxime (a second-generation cephalosporin) over this period of time ([Figure 2](#)). Meanwhile, the use of a more traditional first-generation antibiotic—augmentin—had dramatically declined. In addition, we found that the administration of certain types of antibiotics varied with patients' age ([Figure 3](#)). For instance, cefotaxime (a third-generation cephalosporin) was more commonly used in newborns under 3 months of age, but azithromycin was rarely administered in that age group. This observation might reflect the results of some published studies which reported an increased risk of cardiovascular death associated with the use of azithromycin, specifically, in infants [25,26].

Table 1. Study population demographics.

Characteristics	Year					
	2010	2011	2012	2013	2014	2015
Total population (N)	4216	4062	3935	5080	6198	7021
Sex, n (%)						
Male	2602 (61.72)	2447 (60.24)	2411 (61.27)	3030 (59.65)	3718 (59.99)	4211 (59.98)
Female	1614 (38.28)	1615 (39.76)	1524 (38.73)	2050 (40.35)	2480 (40.01)	2810 (40.02)
Age groups^a, n (%)						
0-3 months	1678 (39.80)	1617 (39.81)	1609 (40.89)	2075 (40.85)	2297 (37.06)	2177 (31.01)
3-6 months	356 (8.44)	413 (10.17)	360 (9.15)	351 (6.91)	410 (6.62)	549 (7.82)
6-12 months	441 (10.46)	491 (12.09)	386 (9.81)	482 (9.49)	580 (9.36)	716 (10.20)
1-2 years	426 (10.10)	421 (10.36)	421 (10.70)	504 (9.92)	723 (11.67)	822 (11.71)
2-5 years	781 (18.52)	725 (17.85)	711 (18.07)	902 (17.76)	1279 (20.64)	1705 (24.28)
>5 years	534 (12.67)	395 (9.72)	448 (11.39)	766 (15.08)	909 (14.67)	1052 (14.98)

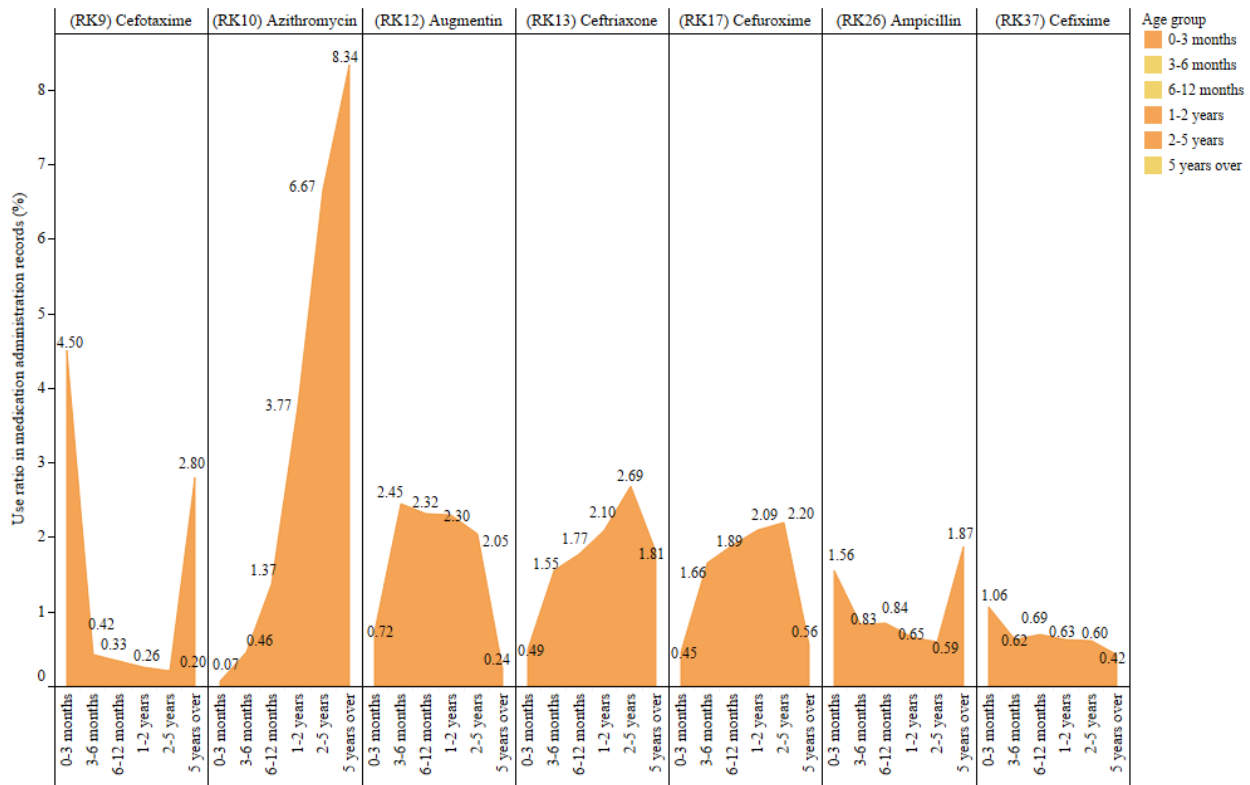
^aAge indicates the age at admission and was calculated as the difference between *Admission Date* and *Birthday*.

Figure 2. The historic proportion of antibiotics administered according to the calendar year.

Using Pattern Mining to Explore the Electronic Medical Record Dataset

We utilized the 3 algorithms to produce the top 10 medication use patterns from 6 age groups spanning from 0 to 17 years, resulting in 180 distinct medication combinations (Multimedia Appendix 1). On the basis of these results, our expert panel summarized 5 clinically interesting medication use patterns as shown in Table 2. Checkmarks indicate that the pattern appeared in that age group. *Pattern 1* is antiasthmatics (albuterol) and expectorants (ipratropium bromide) and corticosteroids

(budesonide). *Pattern 2* is antibiotics and (antiasthmatics or expectorants or corticosteroids). *Pattern 3* is the only use of third-generation antibiotics or the use of these medications followed by traditional antibiotics. *Pattern 4* is antibiotics and medications for enteritis (probiotics: bifid triple viable, smectite, and clostridium butyricum) or medication for skin diseases (antiseptics: zinc oxide and drapolene). Finally, *Pattern 5* is (antiasthmatics or expectorants or corticosteroids) and (medications for enteritis or skin diseases). Detailed descriptions of the patterns for 3 major groups are provided in the following sections.

Figure 3. The historic proportion of antibiotics administered according to patient age group.

Inhaled Medications

Medication use patterns 1 and 2 were both revealed by USpan, because it takes into account a high medication utility for administered medications ([Multimedia Appendix 5](#)). For pattern 1, we found that all 3 medications reflect nonsequential medication administrations, indicating that they were administered concurrently instead of the more traditional sequential application of each medication. Our expert panel confirmed that this medication pattern could shorten medication administration durations, for example by combining and administering multiple inhaled medications simultaneously via nebulization. This approach was also confirmed by the chemical stability of such a mixture [27,28]. Further investigation of the route of administration of pattern 1 medication use indicated that inhalation therapy accounted for 11.53% (78390/680138) of the medication administration records, which we considered as *Treatment pattern 2*. Medication pattern 2 was these inhaled medications plus antibiotics indicating the management pattern of difficult-to-treat infections in patients with pneumonia.

Antibiotics

FP-Growth and PrefixSpan, the 2 classical algorithms, showed a similar pattern 3 ([Table 2](#)), that is, the only use of third-generation antibiotics or the use of these medications with (or followed by) traditional antibiotics. FP-Growth resulted in

2 combinations. One is the combination of 2 third-generation cephalosporins (cefotaxime and cefixime). The other is the combination of a third-generation cephalosporin (ceftriaxone) with a traditional first-generation antibiotic (azithromycin). Interestingly, PrefixSpan showed not only the same 2 medication combinations but also similar timings: the average time gap between the 2 medications was 7 and 5 days, respectively.

Medications for Enteritis or Skin Diseases

Patterns 4 and 5 showed a correlation between pneumonia-specific medications and medications for 2 other conditions: enteritis and skin diseases. For example, pattern 4 demonstrated a third-generation antibiotic (cefotaxime) and a medication for skin rashes (zinc oxide) were followed by a probiotic (bifid triple viable) that appears only in the *0 to 3 months* age group. Pattern 5 showed that the 3 inhalation medications (albuterol, ipratropium bromide, and budesonide) were followed by probiotics without using antibiotics that mainly appeared in the *3 to 6 months* age group. Due to the fact that PrefixSpan only determines the order of different medications without considering a time window, we further examined the time interval between medications for pneumonia following medication for enteritis or skin diseases ([Multimedia Appendix 4](#)). The findings revealed that the average time intervals were similar (4 vs 6 days) in the abovementioned 2 examples for patterns 4 and 5.

Table 2. A select list of clinically interesting results produced by 3 pattern mining algorithms.

#	Medication use patterns	FP-Growth	PrefixSpan	USpan	Age groups					
					0-3 m	3-6 m	6-12 m	1-2 y	2-5 y	>5 y
1	anti-asthmatics AND expectorants AND corticosteroids	(Albuterol, Ipratropium Bromide, Budesonide)	(Albuterol, Ipratropium Bromide, Budesonide)	(Albuterol, Ipratropium Bromide, Budesonide), (Pholcodine)	✓ ^a	✓	✓	✓	✓	✓
2	antibiotics AND (anti-asthmatics OR expectorants OR corticosteroid)	— ^b	—	(Albuterol, Ipratropium Bromide, Budesonide), (Azithromycin)			✓	✓	✓	✓
		—	—	(Albuterol, Ipratropium Bromide, Augmentin, Budesonide), (Azithromycin)						✓
3	third-generation cephalosporin antibiotics with (or followed by) traditional antibiotics	(Cefotaxime, Cefixime)	(Cefotaxime), (Cefixime)	—	✓					
		(Ceftriaxone, Azithromycin)	(Ceftriaxone), (Azithromycin)	—						✓
4	anti-biotics AND (medications for enteritis OR skin diseases)	(Bifid Triple Viable, Cefotaxime, Zinc Oxide)	(Cefotaxime, Zinc Oxide), (Bifid Triple Viable)	—	✓					
		(Bifid Triple Viable, Cefotaxime, Drapolene)	(Bifid Triple Viable, Cefotaxime)	—	✓					
		(Cefotaxime, Zinc Oxide, Drapolene)	(Cefotaxime, Zinc Oxide, Drapolene)	—	✓					
		(Smecitite, Bifid Triple Viable, Cefotaxime)	—	—	✓					
5	(antiasthmatics OR expectorants OR corticosteroids) AND (medications for enteritis OR skin diseases)	(Bifid Triple Viable, Albuterol, Ipratropium Bromide, Budesonide)	(Albuterol, Ipratropium Bromide, Budesonide), (Bifid Triple Viable)	—		✓				

^a✓ indicates that the pattern appeared within the specified age groups.

^bNo result from the specific algorithm.

Discussion

Principal Findings

There is a considerable demand for appropriate and proven approaches to expanding the use of medication analytics. By simply checking the frequency of use of medications in the raw EMR data, we found that IV administration is the most common administration route (treatment pattern 1) for treating inpatient childhood pneumonia in this hospital. This finding is in line with concerns about heavy use of IV therapy in China reported by the WHO [29]. Even then, statistical analysis is limited in its ability to discover relationships among medications, and it typically only produces a list of monotherapies with frequency rankings [30]. Indeed, one could argue that from the perspective of treating a given disease, a medication administered with higher frequency implies higher efficacy. However, our results illustrate that this is not always the case.

By utilizing 3 pattern mining algorithms (FP-Growth, PrefixSpan, and USpan), which are commonly used for business applications, we not only found 5 medication use patterns but also identified an additional treatment pattern (treatment pattern 2), that is, medication administration via inhalation route.

There were differing opinions within our expert panel as to what pattern 4 (antibiotics and medications for enteritis or skin diseases) indicates. One possible explanation, offered by 4 US experts, is that because of the immaturity of infant organs, damage to one organ or organ system could cause reactions in other organs [31]. Furthermore, from a pharmaceutical perspective, broad-spectrum antibiotics (eg, cefotaxime) cause diarrhea by irradiating normal gastrointestinal (GI) flora [32], and physicians can treat this type of diarrhea in the pediatric population with bifid triple viable, a probiotic used to re-establish normal GI flora. Additionally, diarrhea can cause diaper rash that can be treated with zinc oxide. Chinese experts believed that these findings support an underlying internal

relationship between the lungs and large intestine (or the lungs and skin), which is consistent with established theories of traditional Chinese medicine (TCM) [33]. TCM believes that the human body is an organic whole, that there is a relationship between the lung and large intestine, and that the lungs govern the skin and hair [34]. Despite this rationale, our expert panel generally agreed that pattern 5 was unexpected (ie, the use of medications for enteritis or skin diseases following inhalation treatment, without concurrent use of antibiotics). The overall Fleiss kappa for these 7 patterns was 0.693 (Table 2) indicating moderate agreement [35]. Our findings neither contradict nor confirm these hypothesized relationships as our data target treatment protocols which cannot, on their own, speak to the relationship between organ systems. However, according to a recent study on innate lymphoid cells, lung inflammation might originate in the gut [36], and there exists a link between intestinal microbiota and lung diseases (eg, asthma) [37].

Among the strengths of this study is its novel application of the 3 pattern mining algorithms to medical data. We determined 5 medication use patterns that can succinctly reflect prescriber style among complex real-world hospital EMR use. This prescriber style reflects decision factors beyond efficacy (eg, availability and cost of a medication, frequency of administration in a busy setting, and hospital formulary). Although patients with an initial EMR diagnosis of *pneumonia* could include all varieties of pneumonia to which various treatment protocols may apply, most of our results correlated with our literature searches regarding the medications' use. For example, a combination of 2 antibiotics produced by FP-Growth (cefuroxime and azithromycin) was evaluated by Vergis et al [38], who found no increased risk of mortality associated with prescribing the 2 antibiotics continuously or simultaneously. Another example is by Rubio et al [39], which evaluated the sequential combination of cefotaxime and cefixime, finding that prescribing them within 2 to 3 days of each other may result in shorter hospital stays, a pattern that was also found by PrefixSpan. Results that have not been mentioned in the current literature (eg, pattern 3) warrant further investigation.

One limitation of this study is that we focused on one source of data from a single pediatric hospital in Shanghai, China. Although other pediatric hospitals in China might not have unique medical practices, this Shanghai-based pediatric hospital has been the leading medical institution with a strong link to the international pediatric community; it contains up-to-date technology, and its clinicians undergo continuous medical training. Our findings provide firsthand insight into understanding Chinese pediatricians' experiences in the treatment of pneumonia. Although our results may not be generalizable across all demographics, especially in remote areas of China, it is likely that most medication use patterns are universal in the more populated areas of China.

Our work is the first step toward better synthesis of current practice and establishing more realistic treatment protocols for childhood pneumonia. Moving forward, we plan to expand this study within and across childhood pneumonia groups from multiple health care organizations. We also envision that a knowledge base of medication use patterns will serve as an informative guide to researchers and clinicians. Our discovered treatment patterns have the potential for inclusion in treatment protocols. We hope to further integrate clinical and nonclinical information into our algorithms to help determine cost and efficacy of treatments, as well as readmission, mortality, and morbidity rates.

Conclusions

We used a pattern mining approach to automatically acquire knowledge of prior medication treatment combinations for childhood pneumonia. An expert panel summarized 5 medication use patterns. These, together with 2 identified treatment patterns that also targeted skin disease and enteritis, may warrant further investigation. Additionally, our findings suggest the following starting points for further discussion: (1) a comparison of IV therapy before and after the publication of China's new deal on the rational use of medicines, the 2013 *Principle of Rational Use of Medicines* [40], (2) validation and comparison of efficacy of various medication use patterns, and (3) a potential relationship between the lungs and skin.

Acknowledgments

The authors would like to thank the nonauthors in our expert panel: Datian Che, MD, Diane L. Seger, RPh, and Changzheng Yuan, DS, for their time and advice. Also, we would like to thank Siyuan Cheng, MSc, Hai Cao, MSc, Suzanne V. Blackley, MA, and Joseph M Plasek, MS, for helping us improve USpan and revise the manuscript. The work was partially funded by the National Natural Science Foundation of China Project #71473164, 71874110, U1636207 and 91546105, the Shanghai Science and Technology Development Fund #16JC1400801, 17511105502, 17511101702, and the Suzhou Science and Technology Bureau Technology Demonstration Project (SS2017 12, SS201812).

Authors' Contributions

All authors provided substantial contribution to the conception and design of this work, its data analysis and interpretation, and helped draft and revise the manuscript. All the authors are accountable for the integrity of this work. CT, HS, LR, YX, and GY conceived and designed the experiments. CT, HS, VC, LR, YX, and GY analyzed the data. CT, LR, and YX performed the experiments. CT, VC, LR, AA, YX, GY, JM, and DWB contributed reagents, materials, or analysis tools. CT, JY, VC, AA, LR, YX, GY, JM, and DWB wrote and revised the paper. CT, SH, and YX contributed equally to this work. JM and DB are joint senior authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The implementation of the USpan algorithm.

[[PDF File \(Adobe PDF File\), 66KB - medinform_v7i1e12577_app1.pdf](#)]

Multimedia Appendix 2

Related commonly administered medications with their shortened reference names.

[[PDF File \(Adobe PDF File\), 98KB - medinform_v7i1e12577_app2.pdf](#)]

Multimedia Appendix 3

Values for Fleiss kappa (N=7, n=8, k=2) as substantial agreement.

[[PDF File \(Adobe PDF File\), 99KB - medinform_v7i1e12577_app3.pdf](#)]

Multimedia Appendix 4

Appearance of “medications for enteritis or skin diseases” in pneumonia treatment course within patients’ medication administration records.

[[PDF File \(Adobe PDF File\), 66KB - medinform_v7i1e12577_app4.pdf](#)]

Multimedia Appendix 5

The top 10 results distinguished by age groups.

[[PDF File \(Adobe PDF File\), 55KB - medinform_v7i1e12577_app5.pdf](#)]

References

1. World Health Organization. 2018 May 24. The top 10 causes of death URL: <https://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death> [accessed 2019-01-21] [WebCite Cache ID 75ZumyoMe]
2. World Health Organization. 2016 Nov 6. Pneumonia URL: <http://www.who.int/mediacentre/factsheets/fs331/en> [WebCite Cache ID 73rKrnS8q]
3. Guan X, Silk BJ, Li W, Fleischauer AT, Xing X, Jiang X, et al. Pneumonia incidence and mortality in Mainland China: systematic review of Chinese and English literature, 1985-2008. PLoS One 2010 Jul 23;5(7):e11721 [FREE Full text] [doi: [10.1371/journal.pone.0011721](https://doi.org/10.1371/journal.pone.0011721)] [Medline: [20668535](https://pubmed.ncbi.nlm.nih.gov/20668535/)]
4. Currie J, Lin W, Meng J. Addressing antibiotic abuse in China: an experimental audit study. J Dev Econ 2014 Sep 1;110:39-51 [FREE Full text] [doi: [10.1016/j.jdeveco.2014.05.006](https://doi.org/10.1016/j.jdeveco.2014.05.006)] [Medline: [26949284](https://pubmed.ncbi.nlm.nih.gov/26949284/)]
5. Mulholland K. Problems with the WHO guidelines for management of childhood pneumonia. Lancet Glob Health 2018 Dec;6(1):e8-e9 [FREE Full text] [doi: [10.1016/S2214-109X\(17\)30468-0](https://doi.org/10.1016/S2214-109X(17)30468-0)] [Medline: [29241619](https://pubmed.ncbi.nlm.nih.gov/29241619/)]
6. Pritchard JR, Lauffenburger DA, Hemann MT. Understanding resistance to combination chemotherapy. Drug Resist Updat 2012 Oct;15(5-6):249-257 [FREE Full text] [doi: [10.1016/j.drug.2012.10.003](https://doi.org/10.1016/j.drug.2012.10.003)] [Medline: [23164555](https://pubmed.ncbi.nlm.nih.gov/23164555/)]
7. Ostapchuk M, Roberts DM, Haddy R. Community-acquired pneumonia in infants and children. Am Fam Physician 2004 Sep 1;70(5):899-908 [FREE Full text] [Medline: [15368729](https://pubmed.ncbi.nlm.nih.gov/15368729/)]
8. Cordoba G, Siersma V, Lopez-Valcarcel B, Bjerrum L, Llor C, Aabenhus R, et al. Prescribing style and variation in antibiotic prescriptions for sore throat: cross-sectional study across six countries. BMC Fam Pract 2015 Jan 29;16:7 [FREE Full text] [doi: [10.1186/s12875-015-0224-y](https://doi.org/10.1186/s12875-015-0224-y)] [Medline: [25630870](https://pubmed.ncbi.nlm.nih.gov/25630870/)]
9. Wu M, Sirota M, Butte AJ, Chen B. Characteristics of drug combination therapy in oncology by analyzing clinical trial data on ClinicalTrials.gov. Pac Symp Biocomput 2015:68-79 [FREE Full text] [Medline: [25592569](https://pubmed.ncbi.nlm.nih.gov/25592569/)]
10. Chou TC. Drug combination studies and their synergy quantification using the Chou-Talalay method. Cancer Res 2010 Jan 15;70(2):440-446 [FREE Full text] [doi: [10.1158/0008-5472.CAN-09-1947](https://doi.org/10.1158/0008-5472.CAN-09-1947)] [Medline: [20068163](https://pubmed.ncbi.nlm.nih.gov/20068163/)]
11. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013 Jan 1;20(1):144-151 [FREE Full text] [doi: [10.1136/amiajnl-2011-000681](https://doi.org/10.1136/amiajnl-2011-000681)] [Medline: [22733976](https://pubmed.ncbi.nlm.nih.gov/22733976/)]
12. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep 2016 Dec 17;6:26094 [FREE Full text] [doi: [10.1038/srep26094](https://doi.org/10.1038/srep26094)] [Medline: [27185194](https://pubmed.ncbi.nlm.nih.gov/27185194/)]
13. Aggarwal CC, Han J, editors. Frequent Pattern Mining. New York City: Springer; 2014:471.

14. Fournier-Viger P. The Data Mining Blog. 2013 Oct 13. An introduction to frequent pattern mining URL: <http://data-mining.philippe-fournier-viger.com/introduction-frequent-pattern-mining> [accessed 2018-11-12] [WebCite Cache ID 73rL4sloy]
15. Wright AP, Wright AT, McCoy AB, Sittig DF. The use of sequential pattern mining to predict next prescribed medications. *J Biomed Inform* 2015 Feb;53:73-80 [FREE Full text] [doi: [10.1016/j.jbi.2014.09.003](https://doi.org/10.1016/j.jbi.2014.09.003)] [Medline: [25236952](https://pubmed.ncbi.nlm.nih.gov/25236952/)]
16. Last M, Carel R, Barak D. Utilization of data-mining techniques for evaluation of patterns of asthma drugs use by ambulatory patients in a large health maintenance organization. 2007 Presented at: Proceedings of Seventh IEEE International Conference on Data Mining Workshops (ICDMW). IEEE; October 28-31, 2007; Omaha, Nebraska p. 169-174. [doi: [10.1109/ICDMW.2007.50](https://doi.org/10.1109/ICDMW.2007.50)]
17. Wikipedia. 2005. Chinese patent medicine URL: https://en.wikipedia.org/wiki/Chinese_patent_medicine [accessed 2018-11-12] [WebCite Cache ID 73rMgVxTQ]
18. Sun H, Gu Z, Gao C, Yu G. Descriptive analysis of medication mode in hospitalized children with pneumonia by electronic medical records massive data. *Pharm Care Res* 2014 Aug 30;14(4):264-267 [FREE Full text] [doi: [10.5428/pcar20140408](https://doi.org/10.5428/pcar20140408)]
19. Craig T. Pattern builders. 2011 Mar 2. Tales of Beers and Diapers URL: <https://blog.patternbuilders.com/2011/03/02/tales-of-beers-and-diapers> [accessed 2018-11-12] [WebCite Cache ID 73rLN2PrL]
20. Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. 2000 Presented at: SIGMOD '00 Proceedings of the 2000 ACM SIGMOD international conference on Management of data; May 15-18, 2000; Dallas, Texas, USA p. 1-12. [doi: [10.1145/342009.335372](https://doi.org/10.1145/342009.335372)]
21. Pei J, Han J, Mortazavi-Asl B, Wang J, Pinto H, Chen Q, et al. Mining sequential patterns by pattern-growth: the prefixspan approach. *IEEE Trans Knowl Data Eng* 2004 Nov;16(11):1424-1440. [doi: [10.1109/TKDE.2004.77](https://doi.org/10.1109/TKDE.2004.77)]
22. Yin J, Zheng Z, Cao L. USpan: an efficient algorithm for mining high utility sequential patterns. 2012 Presented at: KDD '12 Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; August 12-16, 2012; Beijing, China p. 660-668. [doi: [10.1145/2339530.2339636](https://doi.org/10.1145/2339530.2339636)]
23. Chan R, Yang Q, Shen Y. Mining high utility itemsets. 2003 Presented at: ICDM '03 Proceedings of the Third IEEE International Conference on Data Mining; November 19-22, 2003; Melbourne, Florida p. 19-26. [doi: [10.1109/ICDM.2003.1250893](https://doi.org/10.1109/ICDM.2003.1250893)]
24. Armstrong JS. Findings from evidence-based forecasting: methods for reducing forecast error. *Int J Forecast* 2006 Apr;22(3):583-598. [doi: [10.1016/j.ijforecast.2006.04.006](https://doi.org/10.1016/j.ijforecast.2006.04.006)]
25. Knirsch CA, Chandra R. Azithromycin and the risk of cardiovascular death. *N Engl J Med* 2012 Dec 23;367(8):772-773. [doi: [10.1056/NEJMc1207269](https://doi.org/10.1056/NEJMc1207269)] [Medline: [22913693](https://pubmed.ncbi.nlm.nih.gov/22913693/)]
26. Tilelli JA, Smith KM, Pettignano R. Life-threatening bradyarrhythmia after massive azithromycin overdose. *Pharmacotherapy* 2006 Jan;26(1):147-150. [Medline: [16506357](https://pubmed.ncbi.nlm.nih.gov/16506357/)]
27. McKenzie JE, Cruz-Rivera M. Compatibility of budesonide inhalation suspension with four nebulizing solutions. *Ann Pharmacother* 2004 Jun;38(6):967-972. [doi: [10.1345/aph.1D335](https://doi.org/10.1345/aph.1D335)] [Medline: [15084687](https://pubmed.ncbi.nlm.nih.gov/15084687/)]
28. Melani AS. Effects on aerosol performance of mixing of either budesonide or beclomethasone dipropionate with albuterol and ipratropium bromide. *Respir Care* 2011 Mar;56(3):319-326 [FREE Full text] [doi: [10.4187/respcare.00955](https://doi.org/10.4187/respcare.00955)] [Medline: [21255509](https://pubmed.ncbi.nlm.nih.gov/21255509/)]
29. Kan J, Zhu X, Wang T, Lu R, Spencer PS. Chinese patient demand for intravenous therapy: a preliminary survey. *Lancet* 2015 Oct 1;386:S61. [doi: [10.1016/S0140-6736\(15\)00642-X](https://doi.org/10.1016/S0140-6736(15)00642-X)]
30. Kuhn M, Yates P, Hyde C. In: Zhang L, Kuhn M, Peers I, Altan S, editors. Nonclinical statistics for pharmaceutical and biotechnology industries. New York: Springer; Jan 22, 2016:698.
31. Marx V. Tissue engineering: organs from the lab. *Nature* 2015 Jun 18;522(7556):373-377. [doi: [10.1038/522373a](https://doi.org/10.1038/522373a)] [Medline: [26085275](https://pubmed.ncbi.nlm.nih.gov/26085275/)]
32. Ling Z, Liu X, Cheng Y, Luo Y, Yuan L, Li L, et al. Clostridium butyricum combined with Bifidobacterium infantis probiotic mixture restores fecal microbiota and attenuates systemic inflammation in mice with antibiotic-associated diarrhea. *Biomed Res Int* 2015;2015:582048 [FREE Full text] [doi: [10.1155/2015/582048](https://doi.org/10.1155/2015/582048)] [Medline: [25802855](https://pubmed.ncbi.nlm.nih.gov/25802855/)]
33. Liu P, Wang P, Tian D, Liu J, Chen G, Liu S. Study on traditional Chinese medicine theory of lung being connected with large intestine. *J Tradit Chin Med* 2012 Sep;32(3):482-487 [FREE Full text] [Medline: [23297577](https://pubmed.ncbi.nlm.nih.gov/23297577/)]
34. Yan CL, Zhu YL. The Treatment of External Diseases With Acupuncture and Moxibustion. Boulder, CO: Blue Poppy Pr; Aug 1999:253.
35. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
36. Mjösberg J, Rao A. Lung inflammation originating in the gut. *Science* 2018 Dec 5;359(6371):36-37. [doi: [10.1126/science.aar4301](https://doi.org/10.1126/science.aar4301)] [Medline: [29302003](https://pubmed.ncbi.nlm.nih.gov/29302003/)]
37. Huang Y, Mao K, Chen X, Sun M, Kawabe T, Li W, et al. S1P-dependent interorgan trafficking of group 2 innate lymphoid cells supports host defense. *Science* 2018 Dec 5;359(6371):114-119. [doi: [10.1126/science.aam5809](https://doi.org/10.1126/science.aam5809)] [Medline: [29302015](https://pubmed.ncbi.nlm.nih.gov/29302015/)]
38. Vergis EN, Indorf A, File TM, Phillips J, Bates J, Tan J, et al. Azithromycin vs cefuroxime plus erythromycin for empirical treatment of community-acquired pneumonia in hospitalized patients: a prospective, randomized, multicenter trial. *Arch Intern Med* 2000 May 8;160(9):1294-1300. [Medline: [10809032](https://pubmed.ncbi.nlm.nih.gov/10809032/)]

39. Rubio FG, Cunha CA, Lundgren FL, Lima MP, Teixeira PJ, Oliveira JC, et al. Intravenous azithromycin plus ceftriaxone followed by oral azithromycin for the treatment of inpatients with community-acquired pneumonia: an open-label, non-comparative multicenter trial. *Braz J Infect Dis* 2008 Jun;12(3):202-209 [FREE Full text] [Medline: 18833404]
40. Xiao Y, Zhang J, Zheng B, Zhao L, Li S, Li L. Changes in Chinese policies to promote the rational use of antibiotics. *PLoS Med* 2013 Nov;10(11):e1001556 [FREE Full text] [doi: 10.1371/journal.pmed.1001556] [Medline: 24260030]

Abbreviations

CHS: Children's Hospital of Shanghai

EMR: electronic medical record

GI: gastrointestinal

IV: intravenous

TCM: traditional Chinese medicine

WHO: World Health Organization

Edited by G Eysenbach; submitted 22.10.18; peer-reviewed by Y Chen, MJ Kang, M Wan; comments to author 07.11.18; revised version received 11.11.18; accepted 20.11.18; published 22.03.19.

Please cite as:

Tang C, Sun H, Xiong Y, Yang J, Vitale C, Ruan L, Ai A, Yu G, Ma J, Bates D

Medication Use for Childhood Pneumonia at a Children's Hospital in Shanghai, China: Analysis of Pattern Mining Algorithms

JMIR Med Inform 2019;7(1):e12577

URL: <http://medinform.jmir.org/2019/1/e12577/>

doi: [10.2196/12577](https://doi.org/10.2196/12577)

PMID: [30900998](https://pubmed.ncbi.nlm.nih.gov/30900998/)

©Chunlei Tang, Huajun Sun, Yun Xiong, Jiahong Yang, Christopher Vitale, Lu Ruan, Angela Ai, Guangjun Yu, Jing Ma, David Bates. Originally published in *JMIR Medical Informatics* (<http://medinform.jmir.org>), 22.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Figure Correction: Health Information Technology in Healthcare Quality and Patient Safety: Literature Review

Sue S Feldman^{1*}, RN, MEd, PhD; Scott Buchalter^{2*}, MD; Leslie W Hayes^{3*}, MD

¹Department of Health Services Administration, The University of Alabama at Birmingham, Birmingham, AL, United States

²Pulmonary and Critical Care, The University of Alabama at Birmingham Medical Center, Birmingham, AL, United States

³Department of Pediatrics, The University of Alabama at Birmingham Medical Center, Birmingham, AL, United States

* all authors contributed equally

Corresponding Author:

Sue S Feldman, RN, MEd, PhD

Department of Health Services Administration

The University of Alabama at Birmingham

1716 9th Avenue

SHPB #590K

Birmingham, AL, 35294

United States

Phone: 1 6616188805

Email: suefeldman1009@gmail.com

Related Article:

Correction of: <http://medinform.jmir.org/2018/2/e10264/>

(*JMIR Med Inform* 2019;7(1):e11320) doi:[10.2196/11320](https://doi.org/10.2196/11320)

The authors of “Health Information Technology in Healthcare Quality and Patient Safety: Literature Review” (*JMIR Med Inform* 2018;6(2):e10264) mistakenly provided revised figures that were not identical in every way to the original ones in the paper that they were meant to replace.

During proofreading, the authors were asked to provide updated versions of [Figures 4-7](#) because the versions that were originally provided were not of sufficient resolution. When creating better quality images, the authors accidentally omitted frequencies

(groundedness and densities) that were included in the original images.

The incomplete versions of [Figures 4-7](#) have been replaced and can be viewed below.

The correction will appear in the online version of the paper on the JMIR website on January 3, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Figure 4. ACTION Network Diagram (G=groundedness, D=density).

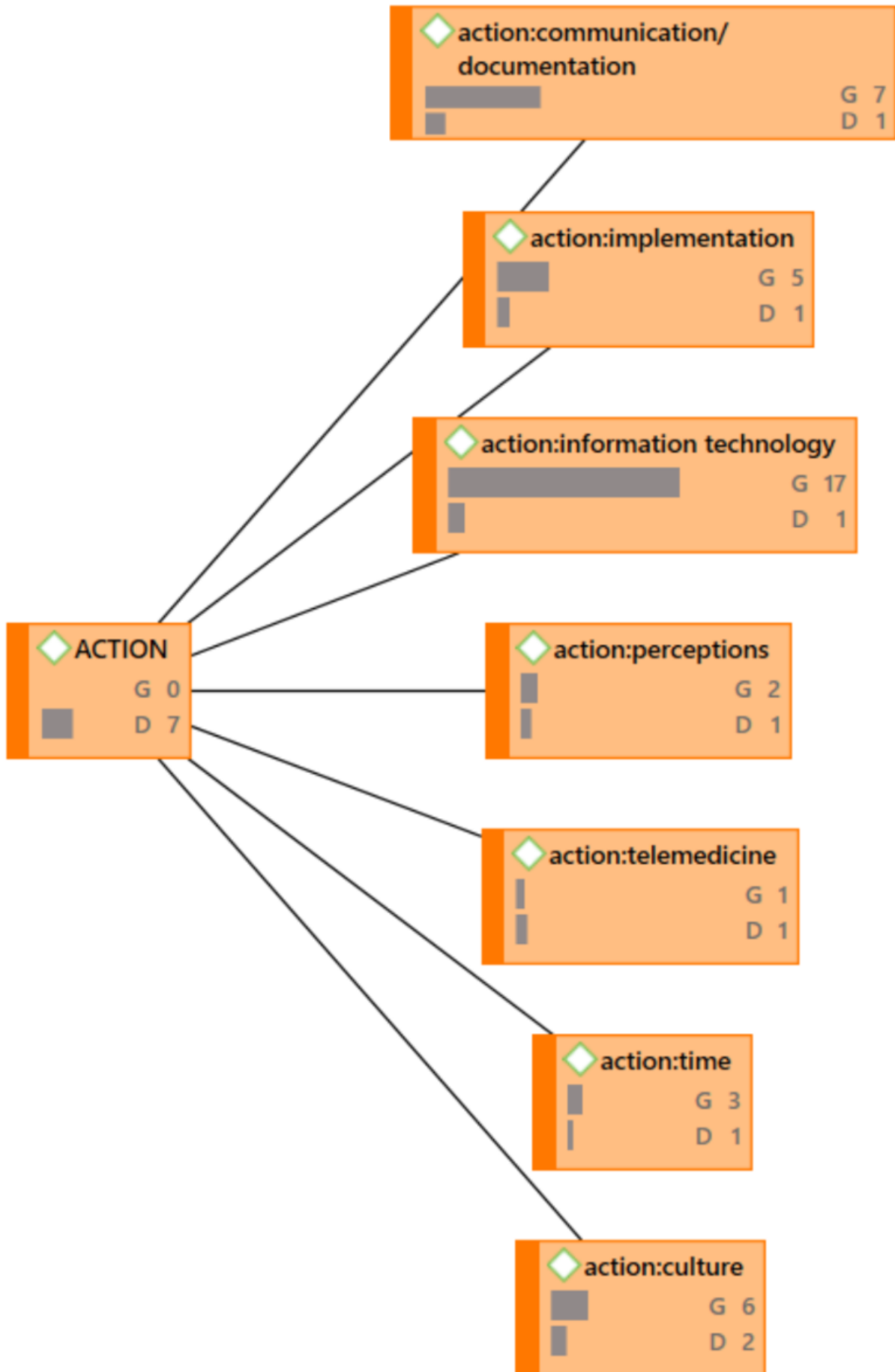


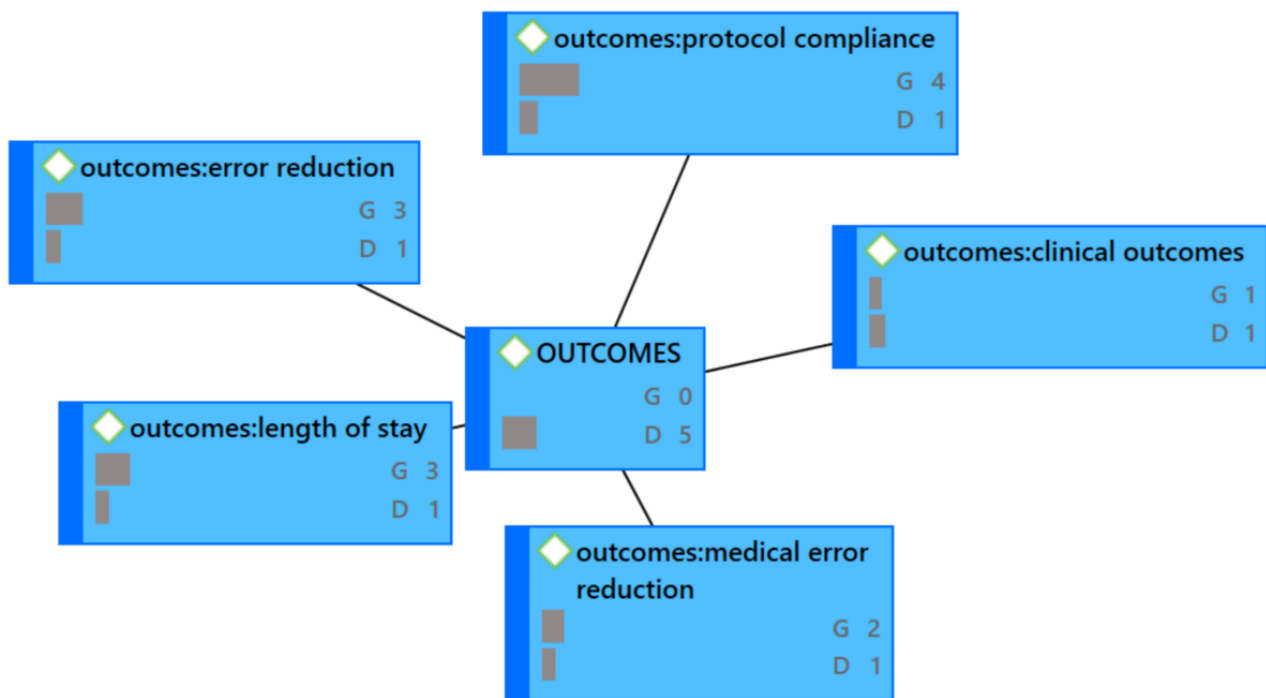
Figure 5. OUTCOMES Network Diagram (G=groundedness, D=density).

Figure 6. IDENTIFICATION Network Diagram (G=groundedness, D=density).

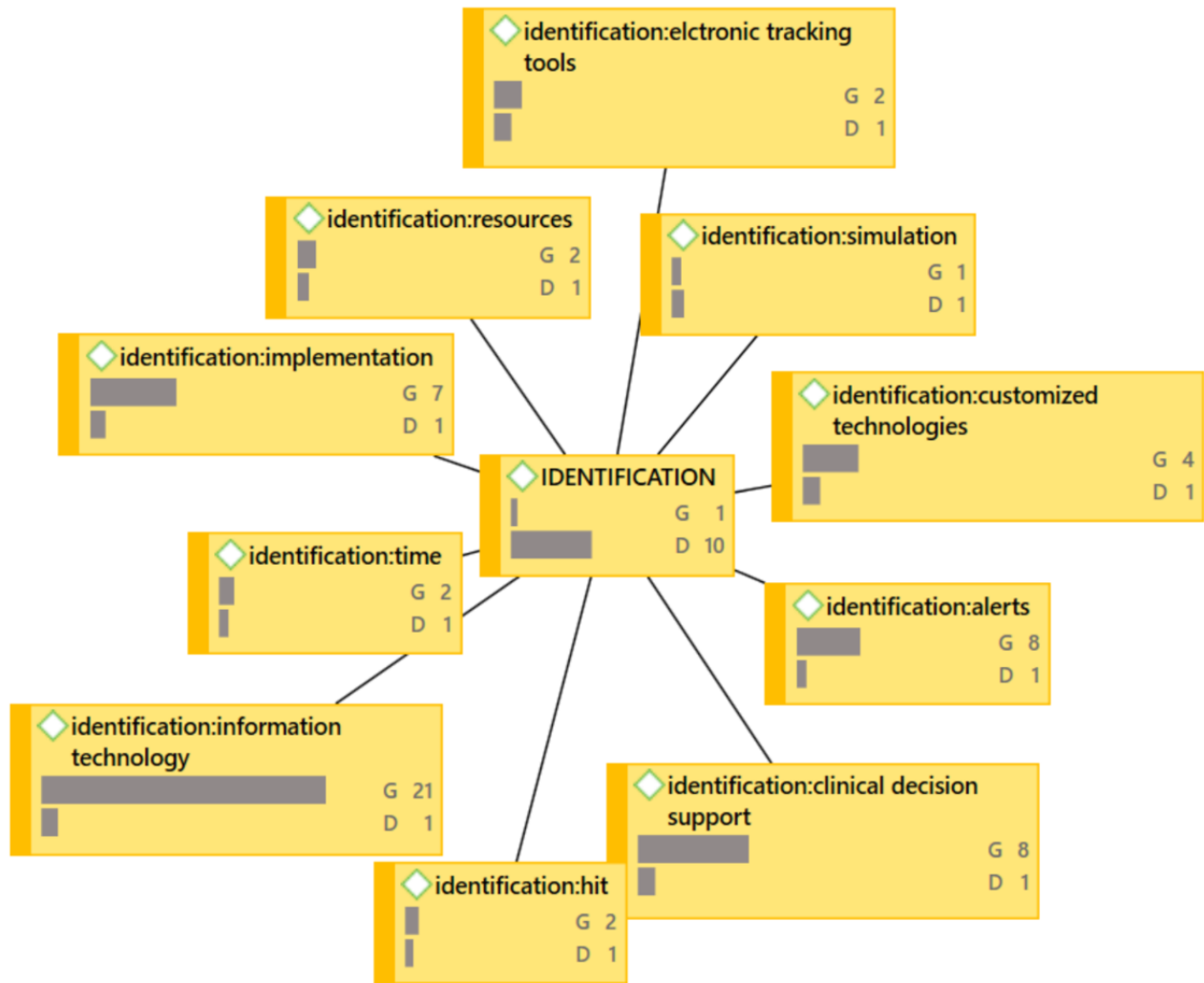


Figure 7. PREVENTION Network Diagram (G=groundedness, D=density).



Edited by G Eysenbach; submitted 18.06.18; this is a non-peer-reviewed article; accepted 09.10.18; published 03.01.19.

Please cite as:

Feldman SS, Buchalter S, Hayes LW

Figure Correction: Health Information Technology in Healthcare Quality and Patient Safety: Literature Review

JMIR Med Inform 2019;7(1):e11320

URL: <https://medinform.jmir.org/2019/1/e11320/>

doi: [10.2196/11320](https://doi.org/10.2196/11320)

PMID: [30609984](https://pubmed.ncbi.nlm.nih.gov/30609984/)

©Sue S Feldman, Scott Buchalter, Leslie W. Hayes. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 03.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>