JMIR Medical Informatics

Impact Factor (2022): 3.2
Volume 3 (2015), Issue 4 ISSN 2291-9694 Editor in Chief: Christian Lovis, MD, MPH, FACMI

Contents

Original Papers

Viewpoint

Short Paper

Original Paper

NHash: Randomized N-Gram Hashing for Distributed Generation of Validatable Unique Study Identifiers in Multicenter Research

Guo-Qiang Zhang[1,2], PhD; Shiqiang Tao[1,2], MS; Guangming Xing[3], PhD; Jeno Mozes[2]; Bilal Zonjy[2,4]; Samden D Lhatoo[2,4], MD, FRCP; Licong Cui[1,2], PhD

[1]Institute of Biomedical Informatics, University of Kentucky, Lexington, KY, United States

[2]Center for SUDEP Research (NINDS-funded Center Without Walls for Collaborative Research in the Epilepsies), Cleveland, OH, United States

[3]Department of Computer Science, Western Kentucky University, Bowling Green, KY, United States

[4]Department of Neurology, Case Western Reserve University, Cleveland, OH, United States

Corresponding Author:
Licong Cui, PhD
Institute of Biomedical Informatics
University of Kentucky
Multidisciplinary Science Building 230
725 Rose Street
Lexington, KY, 40536
United States
Phone: 1 859 218 5936
Fax: 1 859 323 3162
Email: licong.cui@uky.edu

Abstract

Background: A unique study identifier serves as a key for linking research data about a study subject without revealing protected health information in the identifier. While sufficient for single-site and limited-scale studies, the use of common unique study identifiers has several drawbacks for large multicenter studies, where thousands of research participants may be recruited from multiple sites. An important property of study identifiers is error tolerance (or validatable), in that inadvertent editing mistakes during their transmission and use will most likely result in invalid study identifiers.

Objective: This paper introduces a novel method called "Randomized N-gram Hashing (NHash)," for generating unique study identifiers in a distributed and validatable fashion, in multicenter research. NHash has a unique set of properties: (1) it is a pseudonym serving the purpose of linking research data about a study participant for research purposes; (2) it can be generated automatically in a completely distributed fashion with virtually no risk for identifier collision; (3) it incorporates a set of cryptographic hash functions based on N-grams, with a combination of additional encryption techniques such as a shift cipher; (d) it is validatable (error tolerant) in the sense that inadvertent edit errors will mostly result in invalid identifiers.

Methods: NHash consists of 2 phases. First, an intermediate string using randomized N-gram hashing is generated. This string consists of a collection of N-gram hashes f_1, f_2, ..., f_k. The input for each function f_i has 3 components: a random number r, an integer n, and input data m. The result, $f_i(r, n, m)$, is an n-gram of m with a starting position s, which is computed as (r mod |m|), where |m| represents the length of m. The output for Step 1 is the concatenation of the sequence $f_1(r_1, n_1, m_1)$, $f_2(r_2, n_2, m_2)$, ..., $f_k(r_k, n_k, m_k)$. In the second phase, the intermediate string generated in Phase 1 is encrypted using techniques such as shift cipher. The result of the encryption, concatenated with the random number r, is the final NHash study identifier.

Results: We performed experiments using a large synthesized dataset comparing NHash with random strings, and demonstrated neglegible probability for collision. We implemented NHash for the Center for SUDEP Research (CSR), a National Institute for Neurological Disorders and Stroke-funded Center Without Walls for Collaborative Research in the Epilepsies. This multicenter collaboration involves 14 institutions across the United States and Europe, bringing together extensive and diverse expertise to understand sudden unexpected death in epilepsy patients (SUDEP).

Conclusions: The CSR Data Repository has successfully used NHash to link deidentified multimodal clinical data collected in participating CSR institutions, meeting all desired objectives of NHash.

(JMIR Med Inform 2015;3(4):e35) doi:10.2196/medinform.4959

KEYWORDS

cryptographic hash function; multi-center study; study identifiers; health information management; data integration; patient cohort identification; search interface

Introduction

Unique study identifiers, or pseudonyms, are alphanumeric codes used in clinical and other research studies to uniquely identify a study participant without revealing in the identifiers any Personal Health Information (PHI) [1], such as name, full date of birth (DOB), and medical record number (MRN) [2]. For a fictional study participant, Aaron Skotnica, with DOB 08/13/1956 and MRN 07172485, the unique study identifier could be a number such as 57, representing the 57th enrolled study subject. Or it could be a randomly generated number, such as 28262. However, a large number of more sophisticated mechanisms for generating unique study identifiers do exist. A separate codebook, stored and managed in a secure environment, links the unique study identifier to the actual research participant. Electronic data files with unique study identifiers, even de-identified, must be stored in a secure and protected manner, with access granted only to authorized study personnel approved by institutional review boards.

While sufficient for single-site and limited scale studies, the use of common unique study identifiers has several drawbacks for large multicenter studies, where thousands of research participants may be recruited from multiple sites. These drawbacks include (1) the need for coordination so that different sites use distinct blocks of non-overlapping codes for unique study identifier to avoid collision (ie, the same identifier is generated for distinct subjects), (2) difficulty validating if a piece of alphanumeric code is a legitimate unique study identifier or not (eg, if the unique study identifier 57 is inadvertently transposed to 75, the result could be another valid unique study identifier, but for a different study participant), and (3) the possibility that aggregated site-specific information could easily be derived, which may be undesirable if distinct code segments are used for distinct sites, after merging study data in a central study repository.

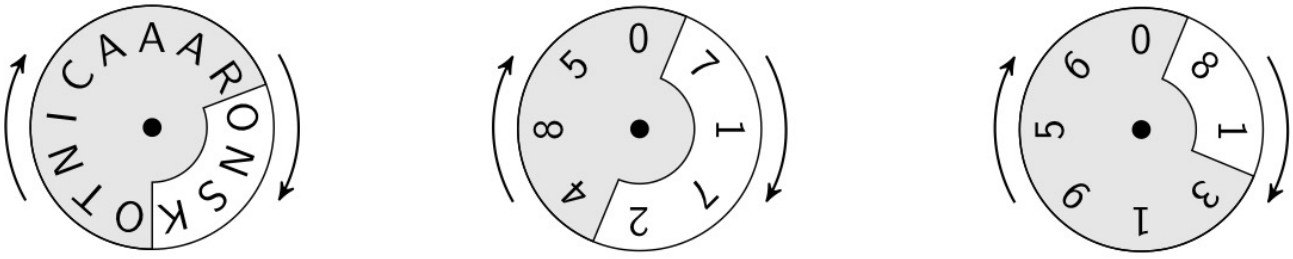
This paper introduces a novel method called Randomized N-gram Hashing (NHash), for generating unique study identifiers for multicenter research. A study identifier generated using NHash has a unique set of properties: (1) as a unique study identifier, it is a pseudonym serving the purpose of linking research data about a study subject for research purposes, (2) it can be generated automatically in a completely distributed and decentralized fashion, yet allowing data integration with virtually no risk for identifier collision, (3) it incorporates a set of cryptographic hash functions based on N-grams for its generation, which can be further encrypted if desired, using encryption techniques such as shift-encryption, and (4) it is validatable in the sense that inadvertent edit errors on NHash identifiers, during their use, will almost always result in invalid identifiers. Furthermore, it is straightforward to validate if the codebook linking study subject and the associated NHash identifier contains errors, simply by regenerating the NHash identifier using the random number with patient information to generate the decryption keys.

For the same fictional study participant, Aaron Skotnica, an NHash identifier is TSXP606170783305. This is achieved by first obtaining an intermediate string, ONSK717281, based on a randomly generated number, 783305. The intermediate string is obtained as the concatenation of the 4-gram of first-name-last-name starting from position x, the 4-gram of MRN starting from position y, and the 2-gram of DOB starting from position z (see Figure 1), where x=3 (783305 mod 13), y=1 (783305 mod 8), and z=1 (783305 mod 8).

The intermediate string, ONSK717281, is then further encrypted to obtain TSXP606170 by shifting each letter in the alphabetic order by 5 (239 mod 26) and each digit by 9 (239 mod 10). The number 239 is calculated using Cantor paring function: $\frac{1}{2}(k_1+k_2)(k_1+k_2+1)+k_2$. Here k_1=13 (length of the participant's full name) and k_2=8 (the participant's birth month). Concatenating TSXP606170 with the random number, 783305, obtains TSXP606170783305 as the final NHash identifier.

We have implemented NHash for the Center for SUDEP Research (CSR), a National Institute for Neurological Disorders and Stroke (NINDS)–funded Center Without Walls for Collaborative Research in the Epilepsies. The CSR Data Repository uses NHash-identifiers to link de-identified multimodal clinical data collected in participating institutions. Since the official launching of CSR in December 2014, 341 study subjects have been enrolled in the Sudden Unexpected Death in Epilepsy Patients (SUDEP) study, with nearly 7TB of EEG signals linked using NHash-identifiers.

Figure 1. An intermediate string, ONSK717281, is generated using the concatenation of 4-gram ONSK from name, 4-gram 7172 from MRN, and 2-gram 81 from DOB, with N-gram hash functions based on a random number 783305.



Background

Protected Health Information

The broader context for effort in the creation of unique study identifiers is the Privacy Rule of the Health Insurance Portability and Accountability Act of 1996 (HIPAA). HIPPA creates a set of requirements and restrictions for the handling of PHI, which is defined as a subset of individually identifiable health information created or received by a health care provider in a variety of forms, including identifiers that could be used to uniquely determine the individual. PHI can come from demographic information, medical history, test and laboratory results, insurance information, and other data that are collected by a health care professional during the delivery of care.

De-identification is a process in which PHI elements are eliminated or manipulated with the purpose of hindering the possibility of revealing PHI contained in the original dataset. This involves removing all identifying data to create unlinkable data. One method of de-identification under HIPPA (called the Safe Harbor Method) used for the current study is when data have been stripped of 18 common identifiers found in patient names, geographic data, all elements of dates, telephone numbers, fax numbers, email addresses, social security numbers, or medical record numbers.

Unique Study Identifiers

A unique study identifier serves the purpose of linking research data about a study subject without revealing PHI in the identifier. With unique study identifiers, de-identified data can be coded with the possibility of linking to the original, fully identified dataset kept by an honest broker or authorized study personnel. The methods for the creation of unique study identifiers range from manual, incremental counts, to completely randomized and encrypted. For example, for the National Database for Autism Research [3], a centralized method for generating global unique identifiers to link collections of research data and specimens is used. For generating such types of global unique identifiers, a Web service is provided for an investigator to input identifying information about a participant into a client application. This information is then encrypted and sent to a server application, returning a generated global unique identifier to the original requester.

Center for SUDEP Research

Epilepsy is the most common serious neurological disorder, affecting 65 million persons worldwide; 200,000 new cases of epilepsy are diagnosed in the United States each year [4]. A third of epilepsy patients fail medical treatment and continue to have seizures [5,6]. SUDEP is the leading mode of epilepsy-related death and is most common in patients with intractable, frequent, and continuing seizures.

The Center for SUDEP Research (CSR) is a National Institute for Neurological Disorders and Stroke (NINDS)–funded Center Without Walls for Collaborative Research in the Epilepsies. This milestone-driven collaboration is composed of researchers from 14 institutions across the United States and Europe and brings together extensive and diverse expertise to SUDEP.

To address the challenges of data integration and data access from multiple Epilepsy Monitoring Units (EMUs), we developed the Multi-Modality Epilepsy Data Capture and Integration System (MEDCIS) [7] that combines retrospective clinical free-text processing using natural language processing (NLP), prospective structured data capture using an ontology-driven interface, and interfaces for cohort search and signal visualization, all in a single integrated environment. A dedicated Epilepsy and Seizure Ontology [8] has been used to streamline the user interfaces, enhance usability, and enable mappings across distributed databases so that federated queries can be executed.

The data capturing component of MEDCIS is called OPIC: Ontology-driven Patient Information Capture [9]. Among the 14 participating institutions, each of the 9 clinical sites will deploy an OPIC instance in their hospital EMUs. A decentralized or distributed study identifier generation method is therefore needed for CSR because (1) there is no centralized management of the codebook to link study identifiers to unique patients, and (2) a global unique identifier generation service involves extra book-keeping and the use of Web services outside the hospital firewall environment.

Methods

Our NHash algorithm generates a unique study identifier for each participant taking, for example, name, MRN, and DOB as input, using randomized N-gram Hashing.

N-Grams

In the fields of computational linguistics, an N-gram is a contiguous sequence of N items from a given sequence of text or speech. The items are basic units of code appropriately defined for each application: it can be syllables, letters, words, or base pairs in bioinformatics. An N-gram of size 1 is referred to as a “unigram,” size 2 is a “bigram,” and size 3 is a “trigram.” For example, ONSK is a 4-gram from AARONSKOTNICA.

Cryptographic Hashing

A cryptographic hash function is a hash function that is considered practically impossible to invert, that is, to recreate the input data from its hash (output) value. These “one-way” hash functions are basic building blocks of modern cryptography. A good cryptographic hash function should have four main properties: (1) it is easy to compute the hash value for any given input data, (2) it is infeasible to generate the input data from merely the hash value, (3) it is infeasible to modify an input data without changing the hash value, and (4) it is infeasible to find two different inputs with the same hash value (false identity).

NHash: Randomized N-gram Hashing for Generating Study Identifiers

NHash consists of two phases. First, an intermediate string using randomized N-gram hashing is generated. This string consists of a collection of N-gram hashes $f_1, f_2, \ldots, f_k$. The input for each function f_i has three components: a random number r, an integer n, and input data m. The result, $f_i(r, n, m)$, is an n-gram of m with a starting position s, which is computed as $(r \bmod |m|)$,

where $|m|$ represents the length of m. The output for Step 1 is the concatenation of the sequence $f_1(r_1, n_1, m_1), f_2(r_2, n_2, m_2), \ldots, f_k(r_k, n_k, m_k)$.

In the second phase, the intermediate string generated in the first phase is encrypted using techniques such as shift-cipher in order not to reveal actual letters in patient names nor digits in DOB or MRN. The result of the encryption, concatenated with the random number r, is the final NHash study identifier.

For the CSR clinical and research data platform, we take k=3, $r=r_1=r_2=r_3$, and m_1 to be the first-name last-name string, m_2 to be the MRN number, and m_3 to be the digital version of DOB. Further, we take n_1=4, n_2=4, and n_3=2, that is, a 4-gram of name, a 4-gram of MRN, and a 2-gram of DOB. The starting positions for these N-grams are determined by r modular the length of m_1, m_2, and m_3, respectively. Table 1 contains an example illustrating the notion of N-gram hash functions based on the fictional study participant, Aaron Skotnica (also shown in Figure 1).

Table 1. Example of the N-gram hash functions used for CSR, generating the final output TSXP606170783305.

i	r_i	n_i	m_i	$\|m_i\|$	$s_i = (r_i \bmod \|m_i\|)$	$f_i(r_i, n_i, m_i)$	Shift-cipher
1	783305	4	AARONSKOTNICA	13	3	ONSK	TSXP
2	783305	4	07172485	8	1	7172	6061
3	783305	2	08131956	8	1	81	70

Figure 2 illustrates the CSR NHash study identifier generation in 6 steps. Given a study participant with required information on name, MRN, and DOB, the input is sanitized first by removing possible punctuation in the name and formatting the date of birth in the format "mmddyyyy." A random number R between 0 and 1,000,000 is generated next. In the third step, a 4-gram from name (name component), a 4-gram of medical record number (MRN component), and a 2-gram from date of birth (DOB component) are extracted from the sanitized input. The starting position for each N-gram is calculated using modular arithmetic. In Step 4, we concatenate these 3 components. In Step 5, we encrypt the concatenated string from Step 4 using such means as shift-cipher. In Step 6, we concatenate the encrypted string and the random number and output it as the unique study identifier. In the same step, we also invoke a duplication check to see if an NHash study ID already exists in the local instance, before the generated study ID is finally put into real use. Step 2 is repeated if a local collision is detected. These six steps are captured more formally as an algorithm in pseudo code (Figure 3).

Figure 2. Illustrative diagram for the algorithm to generate NHash study ID.

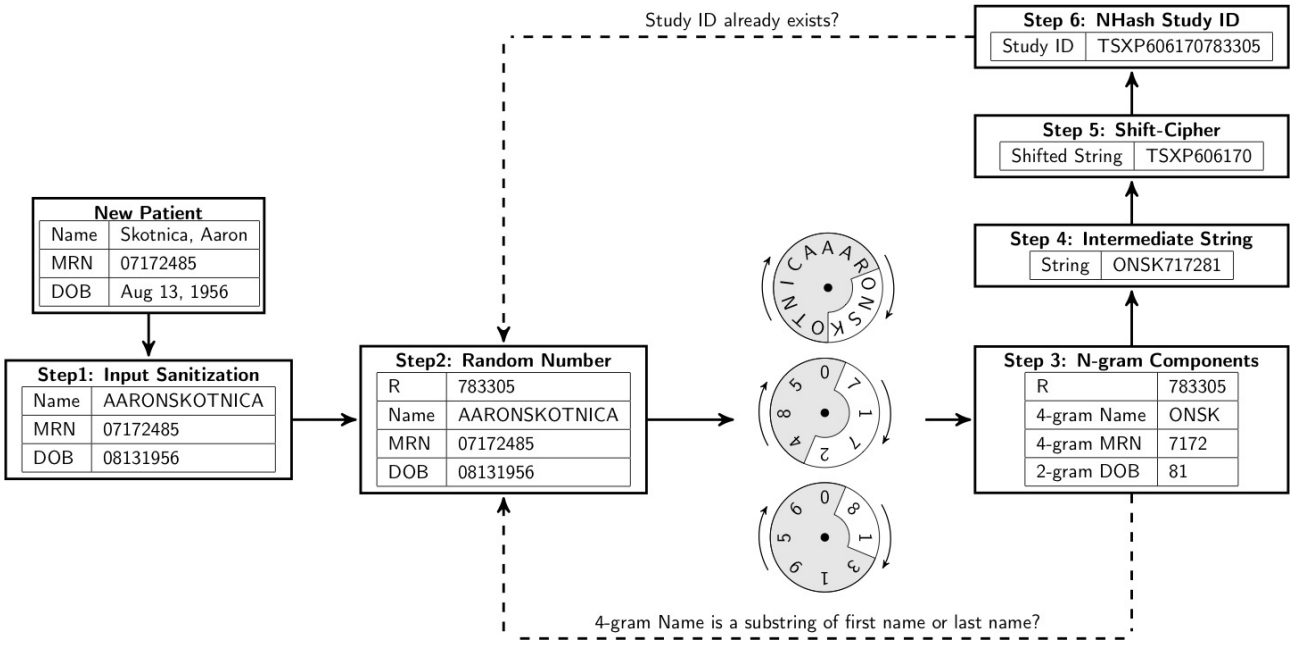


Figure 3. NHash algorithm to generate unique study identifier given patient demographic information.

NHash Identifier Generation

Input: Patient name (Name), medical record number (MRN), and date of birth (DOB)
Output: Unique study identifier (SID) for the given patient

method GENERATE_STUDY_ID(String $Name$, String MRN, Date DOB)
Sanitize Name
Format DOB as "mmddyyyy"
SID = empty string ▷ Initialize an empty string
while SID is empty or SID already exists **do**
    $R = random(0, 1000000)$ ▷ Generate a random number between 0 and 1,000,000
    $start = R \mod Name.length$
    $end = (R + 4) \mod Name.length$
    if $end > start$ **then**
        $Name = Name.substring(start, end)$
    else
        $Name = Name.substring(start, Name.length - 1) + Name.substring(0, end)$
    end if
    $start = R \mod MRN.length$
    $end = (R + 4) \mod MRN.length$
    if $end > start$ **then**
        $MRN = MRN.substring(start, end)$
    else
        $MRN = MRN.substring(start, Name.length - 1) + MRN.substring(0, end)$
    end if
    $start = R \mod DOB.length$
    $end = (R + 2) \mod DOB.length$
    if $end > start$ **then**
        $DOB = DOB.substring(start, end)$
    else
        $DOB = DOB.substring(start, Name.length - 1) + DOB.substring(0, end)$
    end if
    $Str = Name + MRN + DOB$
    $ShiftedStr = shiftCipher(Str, Name, DOB)$ ▷ Shift cipher the concatenated string
    $SID = ShiftedStr + R$
end while
return SID
end method

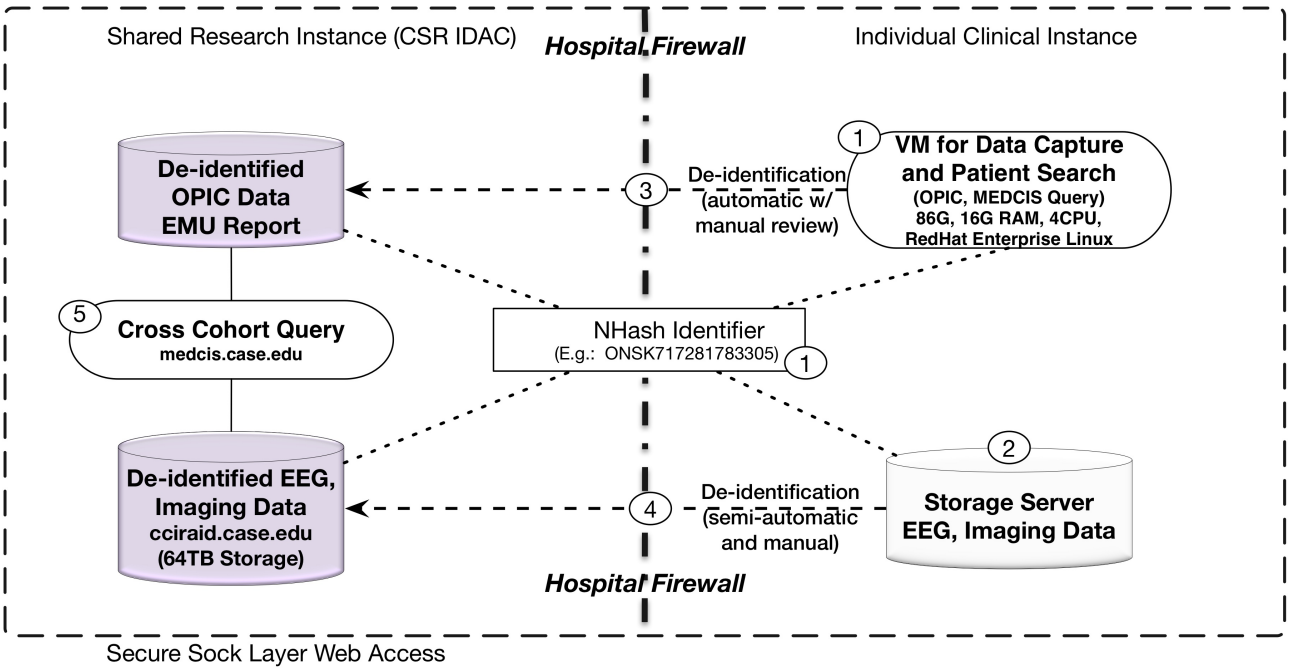
Integration With Ontology-Driven Patient Information Capturing System for Epilepsy

The NHash study identifier generator is an integrated component of OPIC. OPIC captures patient discharge summary reports in EMUs and has a semi-automated built-in de-identification process making clinical data suitable for research for properly enrolled patients. Since one patient can make multiple clinical visits to the EMU and each visit will produce a discharge summary report, the patient study identifier itself is not sufficient enough for naming discharge summary reports directly. To address this phenomenon, we supplement the NHash identifier with an additional 2-digit number and postfix it to the NHash identifier. This way, a discharge summary report is uniquely identified by the NHash identifier plus a 2-digit number postfix. For example, TSXP60617078330501.pdf, TSXP60617078330502.pdf, TSXP60617078330503.pdf, TSXP60617078330504.pdf, if they exist in OPIC, would be the names of discharge summary reports for the first four visits of fictional study subject, Aaron Skotnica.

Linking CSR Multimodal and Multicenter Clinical Research Data

Figure 4 is the CSR data flow architecture for linking multimodal epilepsy research data that includes electrophysiological data in EDF format, imaging data in DICOM format, and genomic data as well. Data flow consists of the following main steps: (1) patient demographics are entered, and unique CSR study identifiers are generated; EMU reports are generated using NHash identifier, (2) larger files (eg, EEG, imaging) are generated from patient care and named using NHash identifier, (3) patient data (demographics, history, medication, diagnosis) and EMU reports are de-identified and exported to the CSR central data repository (IDAC), (4) EEG, imaging, and other larger files are also de-identified and exported to IDAC, and (5) exported data from all CSR clinical sites are aggregated into a single repository (MEDCIS) for searching, querying, and sharing.

Figure 4. CSR data flow and record linking enabled by NHash identifier.

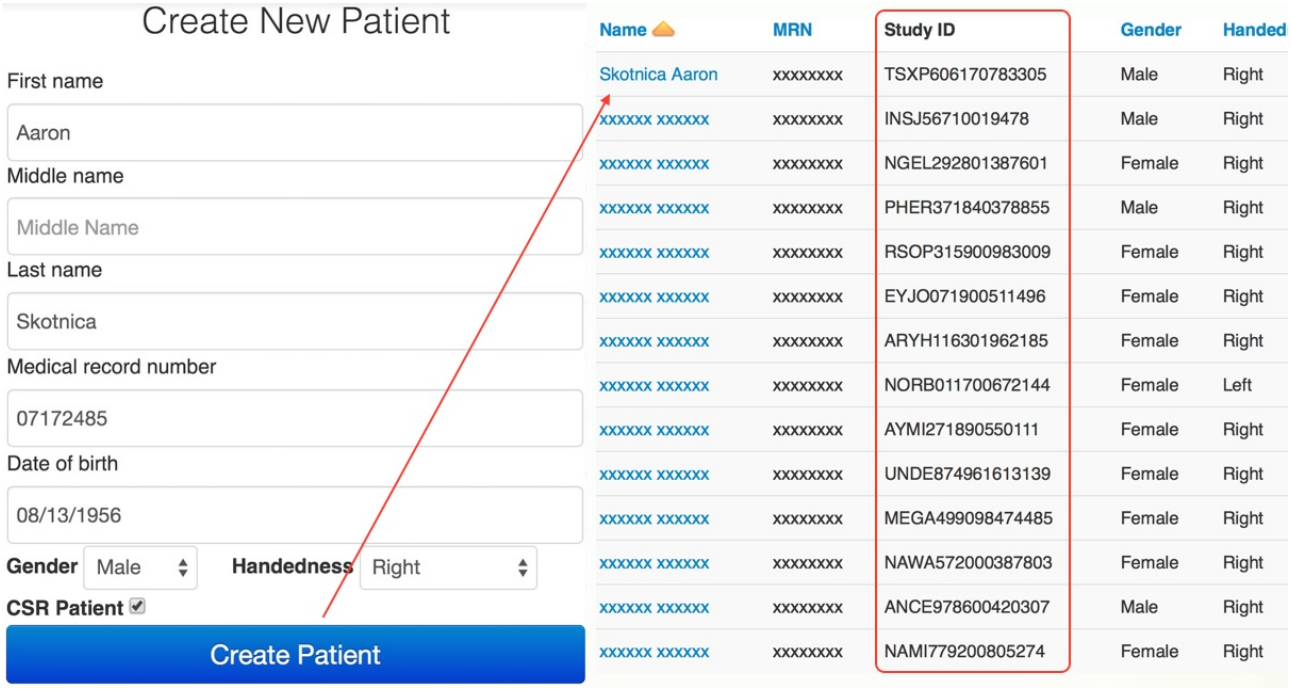


Results

NHash Identifier Generation in Ontology-Driven Patient Information Capture

An NHash identifier is generated at patient enrollment phase using the algorithm in Figure 3 and managed through the Patient Demographics module of OPIC. Figure 5 shows the patient registration interface using our fictional study participant, Aaron Skotnica (left part of Figure 5). For CSR, we use DOB 08/13/1956 and MRN 07172485, among all captured data elements, to generate the unique NHash identifier TSXP606170783305 (top right, Figure 5). Ten additional actual NHash identifiers are displayed inside the area marked by a red rectangle.

Figure 5. Screenshot of OPIC interface generating and displaying NHash identifiers.

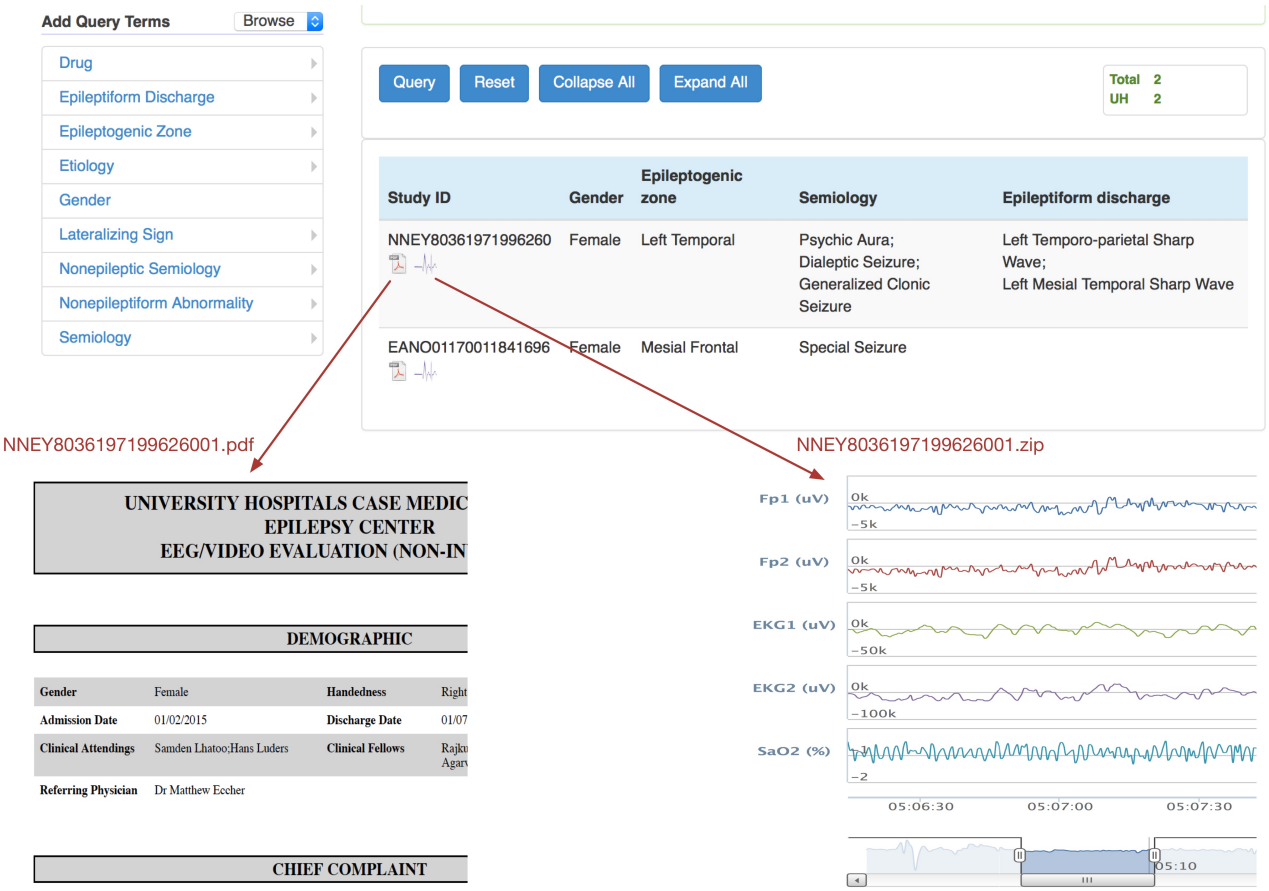


Linking Multimodal Data in MEDCIS

MEDCIS is implemented using agile Web development with the Web application framework, Ruby on Rails. MEDCIS has been deployed [10] for searching, querying, and sharing the CSR multimodal clinical data linked by NHash identifiers. Figure 6 shows a screenshot of the MEDCIS query interface for CSR. The top part is the MEDCIS ontology-driven query interface [7] following the VISAGE design [11]. The results of the query include key characteristics of seizures, as well as links

to related multimodal clinical data, displayed below for discharge summary report in PDF and EEG signals in EDF format. The query result returns the number of study participant reports satisfying the query criteria, as well as detailed participant information including NHash identifier, gender, epileptogenic zone, seizure semiology, and epileptiform discharge. Moreover, MEDCIS provides two hyperlinks (below the NHash identifier) to the de-identified discharge summary report (in PDF) and electrophysiological signal data (in EDF) of the study participant, which can be downloaded or viewed. As indicated in the lower part of Figure 6, NHash identifiers are directly used for linking with associated multimodal clinical data for immediate access.

Figure 6. Screenshot of the MEDCIS query interface for searching, querying, and sharing linked multimodal clinical data.



Deployment

The CSR Institutional Review Board and Data Use Agreements were approved by nine clinical sites involved in the recruitment of patients for the SUDEP study and the deployment of the MEDCIS tools including OPIC, using a common template accessible by participating institutions [12]. Secure Sockets Layer (SSL) has been used for all Web accesses of CSR resources. SSL provides a secure connection between Internet browsers and websites, allowing the transmission of data in encrypted form in transit. Access to the CSR data resource is password-protected, with built-in administrative auditing tools for tracking resource usage activities.

De-identified data were ported to the center repository according to CSR data flow (shown in Figure 4) in two ways: (1) a regular structured database dump to the MEDCIS Web server for cohort search, and (2) secure file transfer protocol (SFTP) to upload larger multimodal files from within hospital firewalls to the data repository. Since the CSR project inception in October 2014, 341 patients have been enrolled (including those enrolled in the pilot phase of the study), and 7TB of data have been moved into the CSR data repository for sharing.

Evaluation

One of the key trade-offs in designing NHash study identifiers is the need for exclusion of any sensitive pieces of data, versus the desire to be able to easily check the validity of a study identifier (even without using the protected codebook). In one extreme of the NHash identifier design, one could use random numbers only, without any of the N-gram hashes (ie, using 0-grams for all components of the identifier). The other extreme is the use of longer N-grams, which would not be desirable because more PHI-related information bits would be revealed.

We believe that a balance could be reached between the two extremes with NHash because our general design of the NHash study identifier is flexible in a number of ways. One is that it allows a variety of study participant information to be included (or not included) for the hash functions. The second is that the sizes of the N-grams (the N) involved are parameters that can be adjusted from study to study, based on specific needs.

Table 2. Analysis of probability for collision.

N-grams	Inverse probability (lower bound)	Inverse probability (upper bound)
Name	10^4	26^4
MRN	9^4	10^4
DOB	10^2	10^2
Random number	10^6	10^6
Total	6.6×10^{15}	4.6×10^{17}
EC for 10^6 records	0.000076	1.09×10^{-6}
EC for 10^7 records	0.0076	1.09×10^{-4}
EC for 10^8 records	0.76	0.0109

The shift operation in the shift-cipher does not change the inverse probability of the ID generation, thus the expected collision is the same with or without the encryption in the last step of NHash. For CSR, we used a 4-gram from name (name component), a 4-gram of medical record number (MRN component), and a 2-gram from date of birth (DOB component), together with a random number between 0 and 1,000,000 to generate the NHash identifier.

One of the basic desired properties of study identifiers is uniqueness: each study participant should have an identifier that is distinct from all other (current and future) identifiers for a study. This uniqueness property is sometimes called collision-free, or free of false identity [3]. Given a hash function with inverse probability of N, the expected number of collisions (EC) after I insertions is $EC=I - N + N(1 - 1/N)^I$.

To see how this formula is derived, let EE be the expected number of empty slots. Then EC=I – (N – EE) since N – EE is the number of occupied slots. For insertion, the probability that a specific slot is occupied is 1/N. Thus the probability of a slot not being occupied by this element is (1 – 1/N). After I insertions, the probability of a slot remaining unoccupied is $(1 – 1/N)^I$. With *N* slots, the expected number of empty slots is $N(1 – 1/N)^I$.

For example, if the inverse probability of the hash function is 6.6×10^{15}(the lower bound inverse probability in Table 2), after 1,000,000 records, the expected number of collisions is $1{,}000{,}000 – 6.6 \times 10^{15} + 6.6 \times 10^{15}(1 – 1/6.6 \times 10^{15})^{1{,}000{,}000}$. This equals 0.000076. In Table 2, we used the lengths of the overwhelming majority of surnames (5) and given names (5) [13,14]. The MRN number is assumed to be 8 digits long, with the possibility of the leading digit being 0. Even though the entire DOB has 8 digits, the year range has only 2 effective digits that can vary in the full range for the current population. Thus, we underestimate the possibility of 2-grams to be 10^2 only. The expected number of collisions also implies that there is only 0.0076% chance for generating a collision after inserting 1,000,000 records. Therefore, for CSR, the probability of collision or false-identity is extremely low, and we have so far encountered no collisions at all in its deployment within a local instance.

Simulation

To further validate the effectiveness of NHash, we performed experiments using a large synthesized dataset comparing NHash with random strings. There are two goals using random string methods in our study: (1) random strings are good benchmark methods for generating unique strings, and (2) validating our experiment design by comparing the expected collision and the number of collisions from the experiment.

It should be noted that the goal of NHash is not to outperform the random string methods in terms of minimizing collision but to provide a hash function with comparable expected collision and the desirable features presented in previous sections.

The synthesized datasets are generated using random names, MRN, and DOB. The names are generated based on the list of top 5000 surnames and 2500 first names in the United States. The frequency of the names is factored in our simulation. The MRN is generated as an 8-digit random number, and the date of birth is a random date from 1910 to 2015. We generated five datasets with 100 million records each.

Table 3 lists the number of collisions using different hash functions. The expected number of collisions is very close to the average number of collisions for all hash functions, and the derivation from the average is low. From Table 3, the number of collisions in NHash-15 is very close to Random-13 and NHash-16 is very close to Random-14, thus clearly indicating the effectiveness of NHash.

Table 3. Number of collisions on synthesized dataset of 100 million (n in Random-n indicates the length of the random string; NHash-15 uses random numbers of length 5 and NHash-16 uses random numbers of length 6).

	Run 1	Run 2	Run 3	Run 4	Run 5	Average	EC
Random -11	1088	1088	1103	1131	1134	1108.8	1086.95
Random-12	110	106	128	123	108	115	108.70
Random-13	9	13	16	8	1	11.2	10.87
Random-14	0	0	0	2	0	0.4	1.09
NHash-15	6	7	5	1	4	4.6	7.58
NHash-16	0	0	1	0	0	0.2	0.76

Discussion

Principal Findings

NHash is a generalizable mechanism for generating study identifiers in a multicenter research setting. In the first phase of generating an intermediate string, the size of randomized N-gram, that is, the number N, is adjustable. In the second phase of further encryption, we used the Cantor paring function for shift-cipher. Other functions and encryption techniques can also be utilized in this phase. The encryption phase provides the mechanism to prevent the NHash scheme from dictionary attacks.

In related work, Johnson [3] proposed a centralized method for generating global unique identifiers to link collections of research data and specimens for Autism spectrum disorder. There are two ways our approach differs. One is that our NHash identifiers are generated distributively, rather than using a centralized Web service. This has advantages in reduced administrative overhead and simplified workflow (since everything is local). Another advantage is that NHash does not require the input of identifiable information into a centralized Web service, which can be perceived as a risk by local sites. The disadvantage is the possibility for false-split: if a study participant is enrolled in two or more sites (on rare occasion), then the participant would be treated as distinct.

The second distinction is in the use of a nondeterministic random number. In Johnson's work [3], the central random identifier generated is the global unique identifier, with a common prefix. For NHash, the random number is part of the identifier but it is also used for generating N-gram hashes that are part of the identifier. This feature makes NHash validatable and much more robust to manual errors compared with typical hash functions. One simple typical method is to assign study participants with sequential unique identifiers according to the creation time. For example, 00001 is assigned for the first study participant, 00002 is assigned for the second study participant, and so forth. Such methods are prone to manual errors, for instance, the study identifier of the first study participant is incorrectly typed as 00002 instead of 00001. In this case, researchers will not be able to tell if the identifier is correct since the study identifier, 00002, by itself represents a valid study participant. However, NHash can easily detect the correctness of the identifier. Taking the same fictional study participant, Aaron Skotnica, as an example. His study identifier is TSXP606170783305. If it is typed as TSXP606170783306 by mistake, NHash is able to detect that it is an invalid study identifier. Based on the algorithm in Figure 3, we know that the postfix 783306 is the random number and that the study identifier would be SXPT06130208783306.

Limitations

The first limitation concerns managing identifier collision in a decentralized setting. Our implementation involves local collision checking (Figure 2) so that all possible identifier collisions within a site are already avoided. We do not have an automated method for checking identifier collision across sites. Even though the probability of identifier collision is extremely small, a process needs to be in place to address it when it happens. We address cross-site identifier collision at data merging stage (see Figure 4). If and when a new batch of data involves an identifier that already exists in the central repository, the local site will regenerate an identifier that is distinct from existing ones. All associated data will be renamed accordingly. However, if the same patient appears at two participating sites, the proposed identifier generation mechanism might not be able to find out that they are the same person.

The second limitation relates to the required data fields to be used for generating NHash. If data for any of the fields are missing or inaccurate, it could create undesirable results. We address this limitation in CSR by reviewing the accuracy and completeness of the patient record before the NHash identifier is generated.

Conclusions

This paper introduces a novel method, NHash, for generating unique study identifiers in a distributed and validatable fashion, in multicenter research involving prospectively collected and de-identified study data. NHash has been deployed for linking multimodal, multicenter data for the Center for SUDEP Research, a National Institute for Neurological Disorders and Stroke–funded Center Without Walls for Collaborative Research in the Epilepsies. Since the official launching of CSR in December 2014, 341 study subjects have enrolled in the SUDEP study, with nearly 7TB of EEG signals linked using identifiers generated using NHash. NHash provides a de-centralized, lightweight study identifier algorithm that offers choices for balancing the trade-offs among the competing requirements of freedom from false-identity and split-identity, and minimal risk for attacks.

Acknowledgments

This research was supported by the National Institute for Neurological Disorders and Stroke (NINDS)–funded Center Without Walls for Collaborative Research in the Epilepsies U01 awards (U01NS090408 and U01NS090405), as well as the University of Kentucky Center for Clinical and Translational Science (Clinical and Translational Science Award UL1TR000117).

Conflicts of Interest

None declared.

References

1. U.S. Department of Health & Human Services. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance PortabilityAccountability Act (HIPAA) Privacy Rule URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html [accessed 2015-11-05] [WebCite Cache ID 6cotjuq9a]
2. Johns Hopkins Bloomberg School of Public Health. Data Security Guidelines for Community-Based Research URL: http://www.jhsph.edu/offices-and-services/institutional-review-board/_pdfs-and-docs/ih-nutrition-gdec-data-security-guidelines-final-2011-05-10.pdf [accessed 2015-11-05] [WebCite Cache ID 6cou9cwNG]
3. Johnson SB, Whitney G, McAuliffe M, Wang H, McCreedy E, Rozenblit L, et al. Using global unique identifiers to link autism collections. J Am Med Inform Assoc 2010;17(6):689-695 [FREE Full text] [doi: 10.1136/jamia.2009.002063] [Medline: 20962132]
4. Epilepsy Foundation. URL: http://www.epilepsy.com/ [accessed 2015-11-05] [WebCite Cache ID 6couJk9nl]
5. Boon P, Vonck K, De HV, Van DA, Goethals M, Goossens L, et al. Deep brain stimulation in patients with refractory temporal lobe epilepsy. Epilepsia 2007 Aug;48(8):1551-1560. [Medline: 17726798]
6. Fisher RS. Emerging antiepileptic drugs. Neurology 1993 Nov;43(11 Suppl 5):S12-S20. [Medline: 8232982]
7. Zhang GQ, Cui L, Lhatoo S, Schuele S, Sahoo S. MEDCIS: Multi-Modality Epilepsy Data Capture and Integration System. AMIA Annu Symp Proc 2014;2014:1248-1257 [FREE Full text] [Medline: 25954436]
8. Sahoo SS, Lhatoo S, Gupta D, Cui L, Zhao M, Jayapandian C, et al. Epilepsy and seizure ontology: towards an epilepsy informatics infrastructure for clinical research and patient care. J Am Med Inform Assoc 2014;21(1):82-89 [FREE Full text] [doi: 10.1136/amiajnl-2013-001696] [Medline: 23686934]
9. Sahoo SS, Zhao M, Luo L, Bozorgi A, Gupta D, Lhatoo SD, et al. OPIC: Ontology-driven Patient Information Capturing system for epilepsy. AMIA Annu Symp Proc 2012;2012:799-808 [FREE Full text] [Medline: 23304354]
10. MEDCIS Cross Cohort Query Interface. URL: https://medcis.case.edu/medcis [accessed 2015-11-05] [WebCite Cache ID 6courzxe7]
11. Zhang GQ, Siegler T, Saxman P, Sandberg N, Mueller R, Johnson N, et al. VISAGE: A Query Interface for Clinical Research. AMIA Jt Summits Transl Sci Proc 2010;2010:76-80 [FREE Full text] [Medline: 21347154]
12. The Center for SUDEP Research (CSR). URL: http://sudepresearch.org [accessed 2015-11-05] [WebCite Cache ID 6cov0REow]
13. The U.S. Census Bureau. URL: http://www.census.gov/topics/population/genealogy/data/2000_surnames.html [accessed 2015-11-05] [WebCite Cache ID 6cov6JmOh]
14. Official Social Security Website. URL: https://www.ssa.gov/OACT/babynames/index.html [accessed 2015-11-05] [WebCite Cache ID 6covOu8EB]

Abbreviations

CSR: Center for SUDEP Research
HIPPA: Privacy Rule of the Health Insurance Portability and Accountability Act
IDAC: Informatics and Data Analytics Core
MEDCIS: Multi-Modality Epilepsy Data Capture and Integration System
NHash: Randomized N-gram Hashing
NINDS: National Institute for Neurological Disorders and Stroke
OPIC: Ontology-Driven Patient Information Capture
PHI: protected health information
SUDEP: sudden unexpected death in epilepsy patients

Edited by G Eysenbach; submitted 20.07.15; peer-reviewed by S Lin; comments to author 11.08.15; accepted 20.09.15; published 10.11.15.

Please cite as:
Zhang GQ, Tao S, Xing G, Mozes J, Zonjy B, Lhatoo SD, Cui L
NHash: Randomized N-Gram Hashing for Distributed Generation of Validatable Unique Study Identifiers in Multicenter Research
JMIR Med Inform 2015;3(4):e35
URL: http://medinform.jmir.org/2015/4/e35/
doi:10.2196/medinform.4959
PMID:26554419

Original Paper

Real-Time and Retrospective Health-Analytics-as-a-Service: A Novel Framework

Hamzeh Khazaei[1], MIEEE, MACM, MEAI, PhD; Carolyn McGregor[2], AM, MACM, SMIEEE, PhD; J Mikael Eklund[3], SMIEEE, PhD; Khalil El-Khatib[2], MIEEE, PhD

[1]IBM, Canada Research and Development Center, Markham, Toronto, ON, Canada

[2]University of Ontario Institute of Technology, Faculty of Business and IT, Oshawa, ON, Canada

[3]University of Ontario Institute of Technology, Department of Electrical, Computer and Software Engineering, Oshawa, ON, Canada

Corresponding Author:
Hamzeh Khazaei, MIEEE, MACM, MEAI, PhD
IBM
Canada Research and Development Center
IBM Canada
3600 Steeles Avenue East
Markham, Toronto, ON, L3R 1H5
Canada
Phone: 1 905 721 8668 ext 3697
Fax: 1 416 567 8167
Email: hamzeh.k.h@ieee.org

Abstract

Background: Analytics-as-a-service (AaaS) is one of the latest provisions emerging from the cloud services family. Utilizing this paradigm of computing in health informatics will benefit patients, care providers, and governments significantly. This work is a novel approach to realize health analytics as services in critical care units in particular.

Objective: To design, implement, evaluate, and deploy an extendable big-data compatible framework for health-analytics-as-a-service that offers both real-time and retrospective analysis.

Methods: We present a novel framework that can realize health data analytics-as-a-service. The framework is flexible and configurable for different scenarios by utilizing the latest technologies and best practices for data acquisition, transformation, storage, analytics, knowledge extraction, and visualization. We have instantiated the proposed method, through the Artemis project, that is, a customization of the framework for live monitoring and retrospective research on premature babies and ill term infants in neonatal intensive care units (NICUs).

Results: We demonstrated the proposed framework in this paper for monitoring NICUs and refer to it as the Artemis-In-Cloud (Artemis-IC) project. A pilot of Artemis has been deployed in the SickKids hospital NICU. By infusing the output of this pilot set up to an analytical model, we predict important performance measures for the final deployment of Artemis-IC. This process can be carried out for other hospitals following the same steps with minimal effort. SickKids' NICU has 36 beds and can classify the patients generally into 5 different types including surgical and premature babies. The arrival rate is estimated as 4.5 patients per day, and the average length of stay was calculated as 16 days. Mean number of medical monitoring algorithms per patient is 9, which renders 311 live algorithms for the whole NICU running on the framework. The memory and computation power required for Artemis-IC to handle the SickKids NICU will be 32 GB and 16 CPU cores, respectively. The required amount of storage was estimated as 8.6 TB per year. There will always be 34.9 patients in SickKids NICU on average. Currently, 46% of patients cannot get admitted to SickKids NICU due to lack of resources. By increasing the capacity to 90 beds, all patients can be accommodated. For such a provisioning, Artemis-IC will need 16 TB of storage per year, 55 GB of memory, and 28 CPU cores.

Conclusions: Our contributions in this work relate to a cloud architecture for the analysis of physiological data for clinical decisions support for tertiary care use. We demonstrate how to size the equipment needed in the cloud for that architecture based on a very realistic assessment of the patient characteristics and the associated clinical decision support algorithms that would be required to run for those patients. We show the principle of how this could be performed and furthermore that it can be replicated for any critical care setting within a tertiary institution.

(JMIR Med Inform 2015;3(4):e36) doi:10.2196/medinform.4640

KEYWORDS

premature babies; physiological data; decision support system; analytics-as-a-service; cloud computing; big data, health informatics; real-time analytics; retrospective analysis; performance modeling

Introduction

Over the past few decades, our society has transitioned to a state where bottlenecks have shifted from a lack of data to limitations in extracting meaningful knowledge from an abundance of data and subsequently using that knowledge to drive decisions. This data-rich, knowledge-poor oxymoron is particularly true in computationally driven clinical decision support systems (CDSSs), where advances in automated high-throughput data acquisition and electronic health records (EHRs) have yet to be translated into knowledge extraction [1].

Adoption of EHRs and systematic collection of physiological data by health care providers were predicted to vastly improve the efficiency and quality of patient care [2]. Unfortunately, despite advances in data collection and storage, these gains have yet to be realized [3,4]. One reason for this failure is that our power to utilize complex, large-scale datasets to generate knowledge and inform clinical decisions remains limited. For example, while CDSSs have existed for decades, they are mostly limited to local alert systems and (data-oblivious) agent-based suggestions that rely on hard-coded criteria.

Recently, enabled by cloud computing Web services, advanced analytics methods have been applied and utilized across a wide spectrum of health care settings for many purposes. Cloud computing has special features for clients (eg, radiologists, physicians, researchers, and patients), aiming to reduce the burden of heavy investments and to utilize resource outsourcing, software, hardware, automated resource management, parallel computing, virtualization, and utility computing [5]. The objectives of such usage include improving patient care, augmenting less-sophisticated rules-based systems, analyzing continuous feeds of physiological data, optimizing financial processes, and resource utilization [6].

Health analytics offers many different methods for the potential improvement of patient care [7]. For example, one predictive risk assessment platform involves using risk assessment analytics to process EHR data to identify patients at the greatest risk for utilizing more resources than their peers with the goal of improving patient outcomes and managing costs. The EHR data were input into a common data model that was then processed by various analytic techniques to stratify patients as "high risk" [8]. Another method described in the literature focused on the potential value of aggregating data enhanced with real-time analytics to provide point-of-care information to oncologists that was tailored to individual patients [9]. One group reported the application of predictive analytics for better targeting of disease management and innovative patient care approaches, while also warning of the unintended consequences that may arise such as excluding disadvantaged populations [10]. Unlabeled and free-text databases such as mammography data can be transformed into computationally accessible collections that are usable for large-scale health analytics [11,12]. Analytics can supplement real-time analysis of physiological data streams in the neonatal intensive care unit (ICU) for earlier detection of worsening medical conditions [13].

Analytics is also utilized in health care applications outside of the traditional inpatient and outpatient patient care settings, such as wearable monitors that patients use at home. Wearable health monitoring systems consist of a variety of sensors, actuators, and multimedia devices, and enable low-cost, noninvasive options for continuous monitoring of health, activity, mobility, and mental status, both indoors and outdoors [14]. Thus, wearable monitoring systems provide continuous physiological data that may reflect the general health of the monitored individuals. The use of wearable sensors in health monitoring systems is an emerging health care field that necessitates data mining and analytics of physiological measurements in a nonclinical setting [15]. Such health monitoring systems may reduce health care costs by disease prevention and enhance the quality of life with disease management and can be tailored to specific uses such as intelligent health monitoring of the elderly individuals in nursing homes and for individuals with dementia or Parkinson's disease [16,17].

These rich sources of data along with aforementioned analytics capabilities have potential for an increased understanding of disease mechanisms and better health care; however, the volume, velocity, variety, veracity, and value of medical data (ie, big data characteristics) present many challenges that limit the effectiveness of outcome for all stakeholders [8]. One promising solution that addresses all these barriers is the Health-Analytics-as-a-Service (HAaaS) paradigm. Analytics-as-a-service (AaaS), in general, is a new "as-a-service," and it is more than just simplifying access to technology. AaaS combines the on-demand aspects of cloud computing with the democratization of information enabled by big data analytics.

In this paper, we present and evaluate a cloud-based reference framework for providing HAaaS for both real-time and retrospective analysis. The framework has the capability to provide all 4 types of analytics, that is, descriptive, predictive, prescriptive, and discovery [18], in a service-oriented fashion. It leverages the latest technologies and best practices for big data analytics and also utilizes the security and privacy measures appropriate for health and medical data. The architecture has been realized (or customized) for neonatal intensive care units (NICUs) at The Hospital for Sick Children (SickKids Hospital) in Toronto and is known as the Artemis project. We have also developed an analytical model for evaluating the performance and availability of an Artemis-IC platform in preparation for migrating Artemis to Artemis-IC. We discuss the important aspects of the system performance and capacity planning process. The main functionalities of the framework are presented

via one of our developed algorithms (ie, Sepsis disease detection). We also present a high-level security and privacy schema for the framework that can be customized and extended for different health applications and use cases. We show the principle of how this could be performed and show that it can be replicated for any critical care setting within a tertiary institution that has critical care.

Methods

In this section, we highlight the functional and nonfunctional characteristics of the framework. Two editions of the framework, research and clinical editions, are designed in such a way that support acquisition and storage of physiological data as well as clinical information, for example, EHR, for the purpose of real-time/retrospective analytics and visualization. The framework is capable of gathering physiological data from a vast variety of medical devices and transfers them in a secure way toward the back-end system residing on the cloud. However, anonymization and potential translation are in order before data leave the hospitals.

The framework has an interface for communication with each hospital's clinical information management system to obtain complementary information (eg, admission information, laboratory test results) of patients. The framework utilizes a hospital interface, which facilitates the management of hospitals' connectivity in various geographic locations. A hospital interface can also be used for "extract, transform, and load" (ETL) purposes as well as load balancing.

Even though the research edition is for retrospective analysis and historic data visualization, it is capable of medical rule deployment and real-time analytics. This is only for testing the new and modified medical rules before undergoing further assessment and auditing. By contrast, the clinical edition was specifically designed for real-time monitoring/visualization, and here human domain experts deploy new or modified medical rules after being extensively validated and certified.

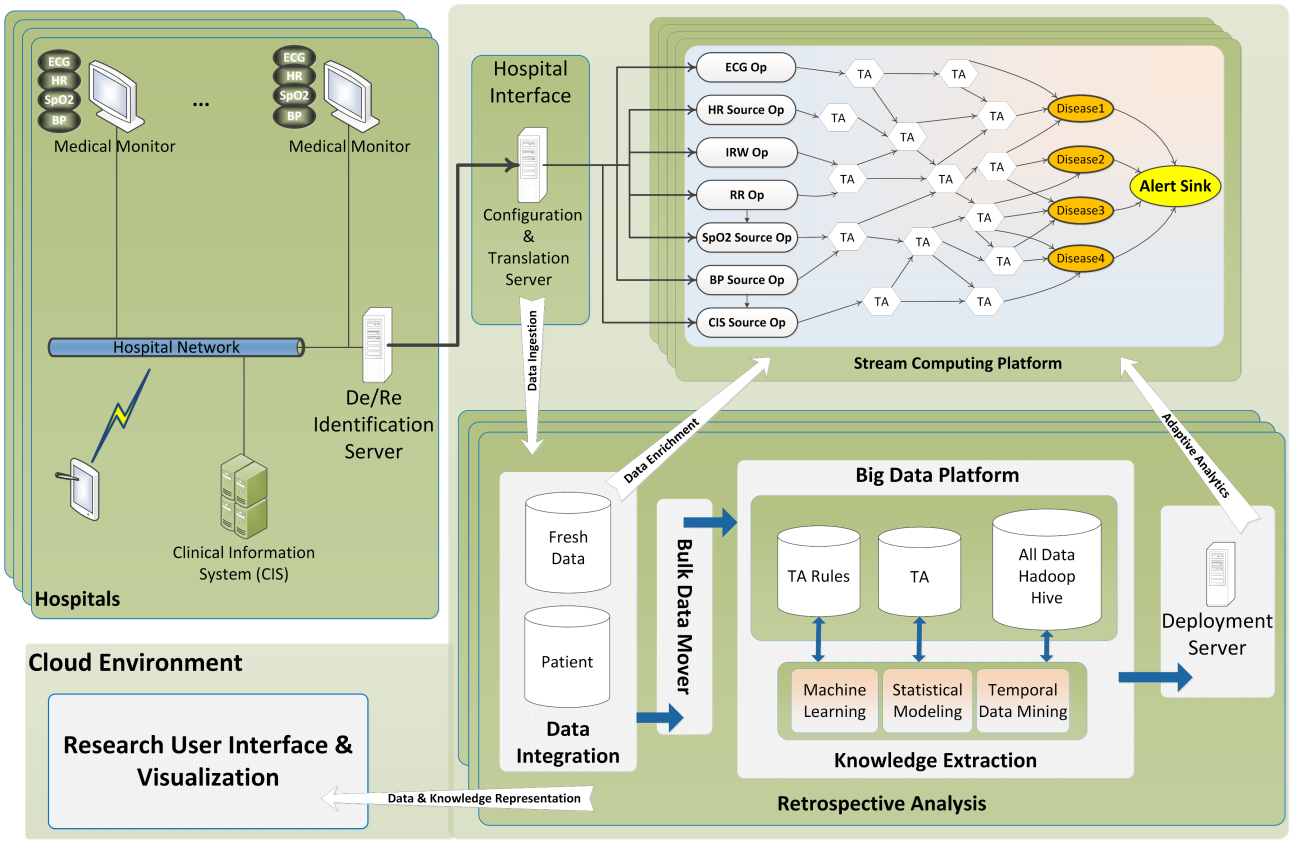
Research Edition

Researchers are the main users of the research edition (RE). This edition can be considered as a comprehensive solution that facilitates retrospective analysis on large numbers of patient data from different places. In addition to real-time analytics capabilities, the RE is able to provide at-rest analytics for stored data. Incorporating a big data analytics solution, that is, Apache Hadoop, offers great power of analysis as well as persistent storage. More specifically, the RE provides clean and ready-to-process medical data (ie, physiological, medical, laboratory, and other complementary data) along with the tools from the Hadoop ecosystem for the researchers to perform their analytics much easier than in the past. Researchers may apply knowledge discovery techniques, for example, temporal data mining [13], machine learning, and statistical modeling, against vast amounts of stored data and find new rules that may help earlier detection of diseases. Such new rules or modified parameters can be deployed to the real-time analysis framework seamlessly. As can be seen in Figure 1, four distinct processes can be identified in the research edition framework.

1. *Data Ingestion:* A process that makes sure that RE stores all relevant data in the Hadoop-based platform.
2. *Data Enrichment:* Historical context that is generated from the data analytics component to bootstrap analytics and enrich incoming data on real-time processing component; more specifically, patient medical data or other related persistent data to enrich the live physiological data during the online processing.
3. *Adaptive Analytics:* Models that are generated by analytics such as data mining, machine learning, or statistical modeling in Hadoop platform used as basis for analytics on incoming physiological data in the real-time component and updated based on online observations.
4. *Data Visualization:* A process that visualizes data and information for different types of users.

In the "Sepsis Case Study" section, we elaborate the data flow and processing steps of the RE in which we describe one of our developed algorithms for detecting sepsis in neonates.

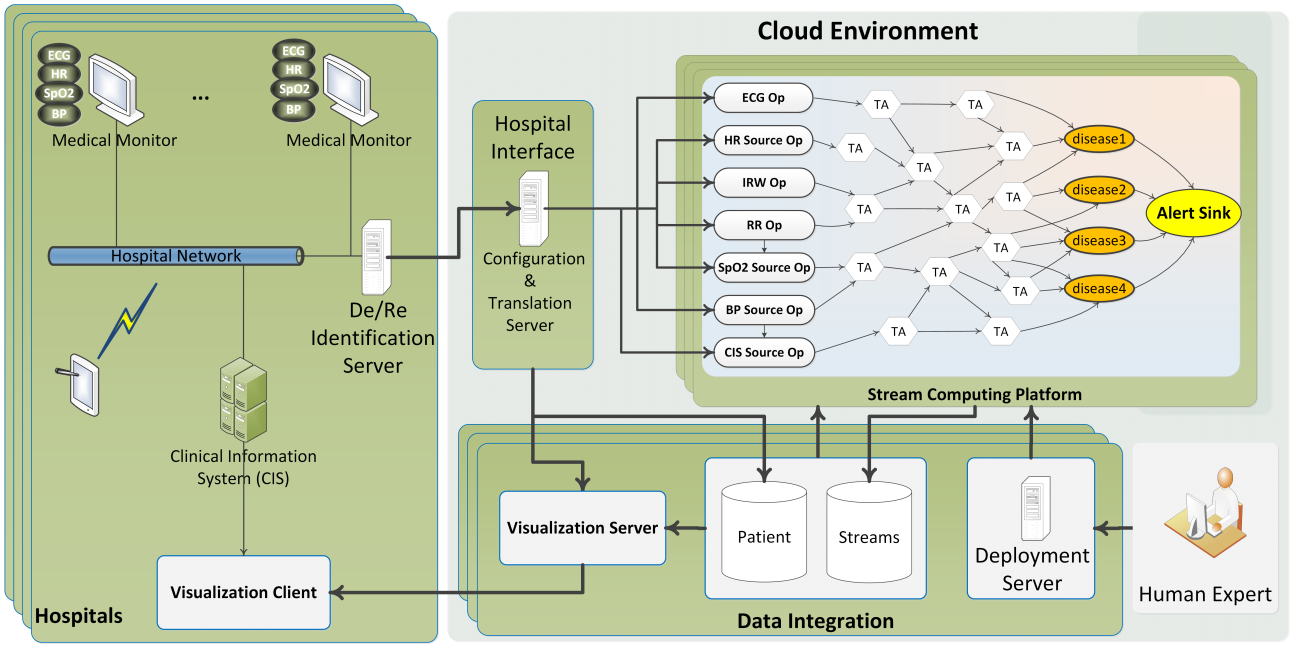
Figure 1. General architecture of the framework (research edition).



Clinical Edition

Clinicians, nurses, specialists, and other authorized hospital staff may use the clinical edition (CE; see Figure 2) to monitor their patients in a much more effective manner in real time. The CE can be considered as a CDSS that can continuously monitor a large number of patients simultaneously and automatically. This edition is capable of monitoring large numbers of patients' physiological/clinical data and producing appropriate alarms in case of any medical complication onset. In addition, it can visualize a specific patient's data either live or historically back a week or more. The ontology for the collection of high-speed synchronous physiological data provides a standardized terminology for acquired physiological data, including measurement metrics, sampling frequency, and acceptable ranges for the received values [19]. As with the collection of physiological data, asynchronous clinical data collection is supported by an ontology that specifies acceptable ranges for the collected values. Examples of clinical data include age, gender, medical history, and laboratory results. The core of the CE is a stream computing middleware component, which provides scalable processing of multiple streams of high-volume, high-rate data.

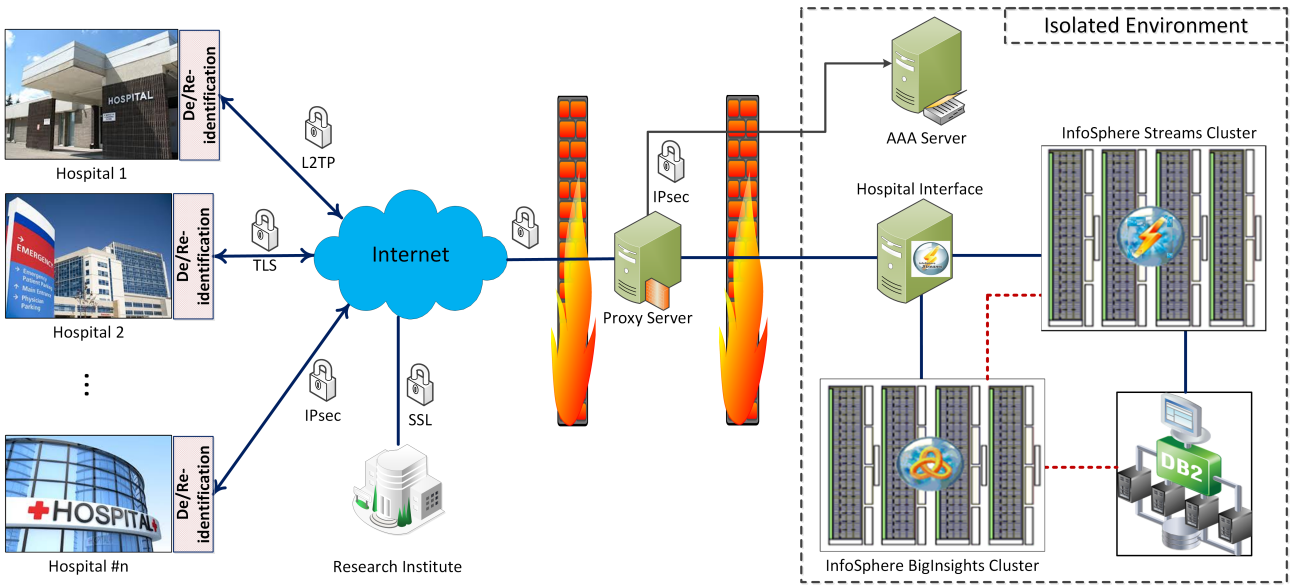
Figure 2. General architecture of the framework (clinical edition).



High-Level Security and Privacy Schema

In this section, we present a high-level security architectural view of the framework. The details and implementation could vary depending on circumstances and applications. As can be seen in Figure 3, hospitals and research institutes are connected to the framework back end through secure channels. Two firewalls have been designed to isolate the framework from the outside world sequentially. The outer one separates the proxy server (ie, framework gateway), which is the edge server of the framework from the Internet. The inner firewall isolates the core of framework from the proxy server. Depending on the granularity of health analytics services, different type of users with various permission and data access levels could be defined. In Artemis-IC, we used a deidentification technique by which we eliminate the properties that might be used to identify patients. Personal data such as medical record number (MRN), name, address, and exact birth date were removed. The MRN was replaced with a study identifier with the translation between the two known only within the hospital. The exact date of birth was replaced with an admission age range of the form 0-3 days old, 4-7 days old, 8-10 days old, and greater than 10 days old. These ranges were chosen for clinically significant reasons. This process is performed in the De/Reidentification Server at hospitals (Figure 2).

Figure 3. Security and privacy perspective of the Artemis-IC framework.



Tailoring of the Method for Monitoring Premature Babies

Premature birth, also known as "preterm birth," is defined as birth before 37 weeks' gestational age. It has been identified as one of the most important perinatal health problems in industrialized nations. NICUs internationally provide critical care for premature and ill term infants. Premature infants in NICUs can be as young as 23 weeks' gestation [20].

Vital organ monitoring together with ventilation support and nutrition or drug titration through smart infusion pumps all generate large volumes of data at high frequency. An electrocardiogram (ECG) graph can be generated based on 1000 readings a second. Heart rate, respiration rate, and blood oxygen are displayed each second resulting in 86,400 readings each day. A premature newborn infant's heart beats more than 7000 times an hour, which is approximately 170,000 times a day. Yet traditional charting protocols, whether documented on paper or within an EHR, typically enable the persistent storage of one value per hour of an indicative heart rate for that hour. A newborn infant's neurological function could also be monitored resulting in multiple waveforms each generating tens of millions of data points per patient per day. Drug and nutrition infusion data from smart infusion pumps can be more than 60 different fields provided every 10 seconds. Given that these infants can have more than 10 infusions concurrently, infusion can generate more than 1 GB of drug infusion data from a single patient per day [21].

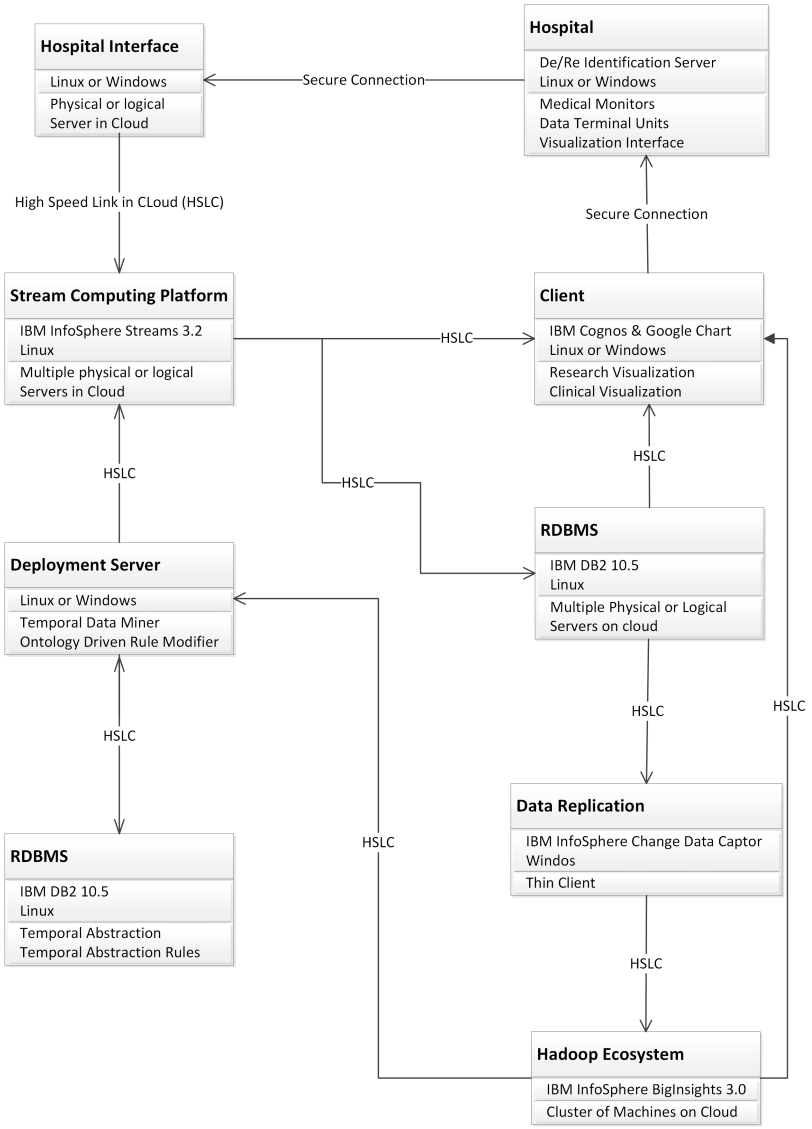
We propose a customized version of the framework, Artemis-IC, for monitoring preterm/surgical babies at NICUs. The Artemis-IC provides HAaaS for concurrent multipatient, multistream, and multidiagnosis through temporal analysis to support real-time clinical decision support and clinical research [22,23]. We deployed a pilot project by implementing Artemis-IC at Toronto's SickKids hospital and proposed an analytical model [24] to enable performance evaluation and capacity planning in advance of final deployment. In addition, there is another pilot of the Artemis-IC at Women and Infants Hospital of Rhode Island (WIHRI), which is collecting physiological data for analytical and simulation modeling purposes. Figure 4 shows the customization and tools that we employed to deploy Artemis-IC framework in SickKids Hospital. As IBM is one of the partners in this research, we used IBM products to implement the framework.

To date, these environments (ie, SickKids and WIHRI deployments) support clinical studies on late-onset neonatal sepsis [22,25]; apnea of prematurity, in which the infant experiences pauses in breathing and reductions in heart rate and blood oxygen saturation [26]; retinopathy of prematurity, which can result in permanent blindness [27]; and pain [28].

Clinicians and researchers are leading these studies from different institutes toward the certification and formal approval of the medical algorithms. Algorithms for the Artemis-IC platform are developed either using data mining techniques that have not previously been detectable, such as our work on late-onset neonatal sepsis [22,25] or identifying patterns described in the medical literature using automated methods such as our work on apnea of prematurity [26]. These algorithms are validated in robust clinical trials before being used to provide decision support for clinicians. For example, the clinical rule states that "If a pause in breathing occurs for greater than 20 seconds, or a pause in breathing that is associated with a change in heart rate, or blood oxygen saturations happens," then a reportable condition of apnea is present [26].

The current Artemis-IC implementations at SickKids and WIHRI have no impact on bedside care, as yet. We are comparing analytical results with current clinical observation and treatment practices to discover new patterns in real-time physiological data that could lead to the earlier detection and prevention of various diseases [26]. From first quarter 2015, we plan to deploy new research where we will be able to compare the results of using Artemis-IC with clinical outcomes using current clinical practices. Some of the algorithms that we have validated when they were running in parallel are due to be certified in 2015/2016 and will be deployed in target clinical institutions. We plan to provide experimental evaluation from multiple deployments of the Artemis-IC in our future reports.

Figure 4. Artemis deployment at SickKids Hospital.

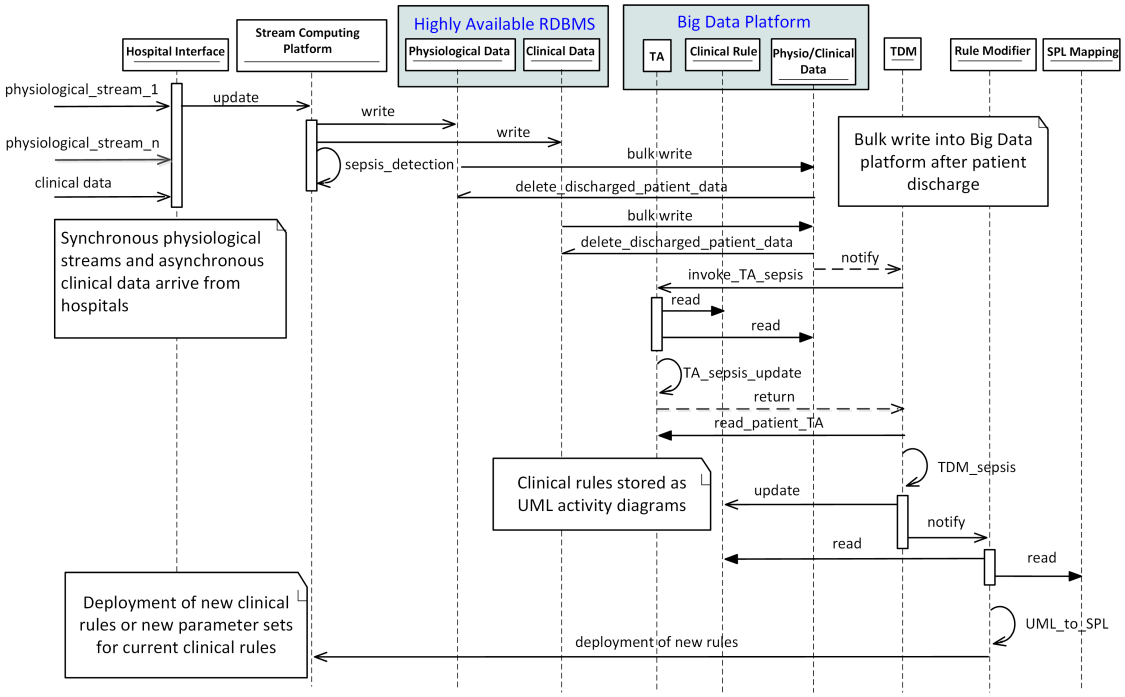


Sepsis Case Study

In this section, we elaborate the interactions between the main components of Artemis-IC for sepsis detection. Sepsis is a potentially life-threatening complication of an infection, which causes whole-body inflammation. In addition to real-time detection, we also demonstrate the knowledge extraction process in detail. The Unified Modeling Language (UML) sequence diagram shown in Figure 5 illustrates all steps including data acquisition, online detection, temporary data storage, persistent data storage within the big data platform, knowledge discovery, knowledge translation, and rule deployment.

Initially, multiple concurrent physiological data streams along with related clinical data are received by the hospital interface. Data are sent to the physiological and clinical database via the stream-computing platform. At the same time, the stream-computing platform runs the current deployed medical rule for sepsis detection. Upon patient discharge, their data including physiological and clinical data will be loaded into the big data platform by the relational database management system (ie, bulk move). Temporal abstractions (TAs) are then performed for the specific service of critical care, in this case sepsis detection, which involves (1) reading from the clinical rules and physiological/clinical tables, and (2) writing the patient TA to the TA table. Temporal data mining then can be performed on the TA results, possibly resulting in updates to the clinical rule table, after null hypothesis-based testing or other rule assessment, for example. Note that the resulting clinical rules are modeled in a UML concurrent activity diagram [19]. The rule modifier is notified of a rule modification and translates the UML representations of the new clinical rule to stream processing language (SPL) based on the SPL mappings active ontology. Finally, the new rule can be deployed on the stream-computing platform for upcoming real-time analysis. Note that the rule deployment on the Artemis-IC clinical edition will be performed under supervision of domain human experts as opposed to here where we consider the Artemis-IC research edition.

Figure 5. UML sequence diagram of sepsis detection and temporal data mining steps.



Quality of Service

As the framework has a service-oriented architecture (SOA), the quality of service (QoS) is of great importance. To assign the proper amount of resources to each hospital, we present a method to create an analytical model to enable an accurate estimation of storage, memory, and computation power for the real-time health analytics components and retrospective analytics components. The model utilizes realistic patient population distribution that is based on gestation age characteristics and condition onset probabilities within those contexts. Both of these variables dictate the predicted length of stay for that infant. In the following section, we present the model within the context of SickKids hospital. In future work, we will do this for other hospitals before deployment. We also leave performance modeling of the research edition as our future work in which we concentrate on another type of users of the framework (ie, researchers).

Analytical Modeling of the Method

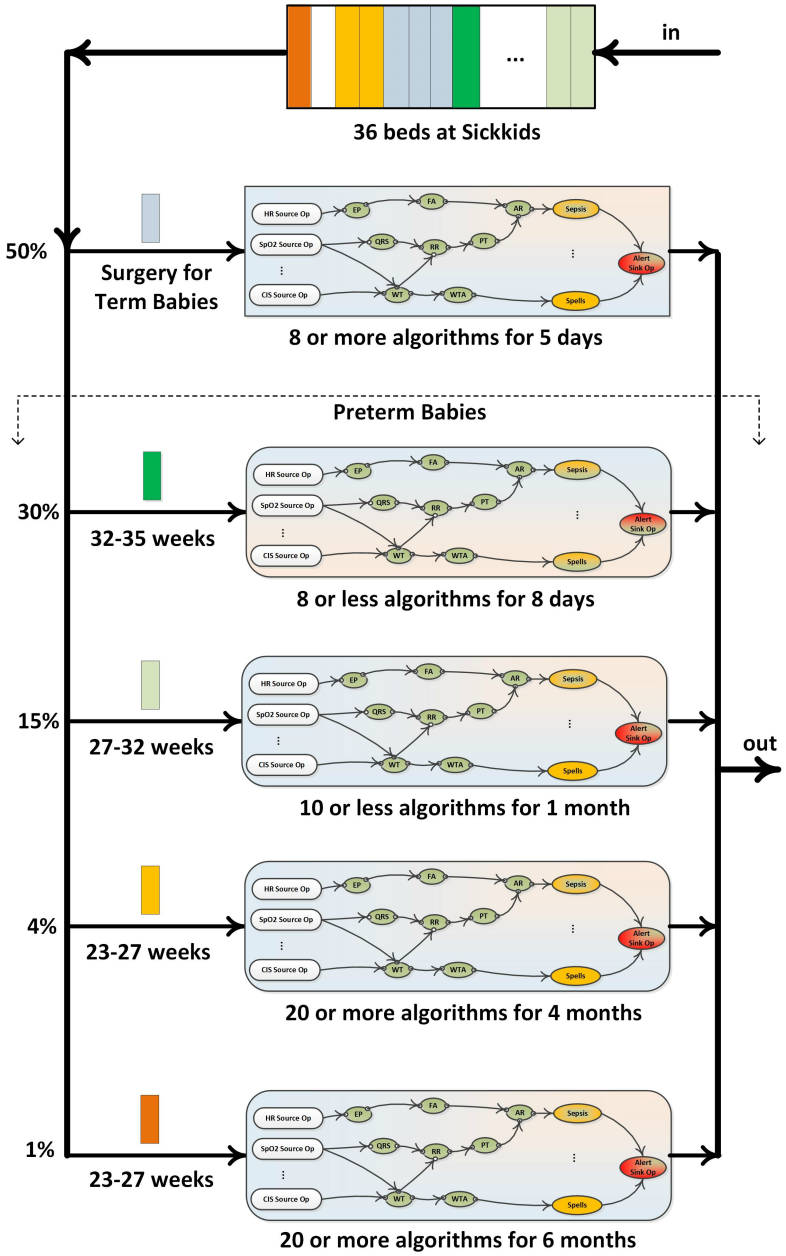
The analytical modeling of Artemis-IC deployment at SickKids hospital's NICU is required before any deployment because critical care units (CCUs)/ICUs are different in terms of types of patients, arrival process of patients, mean hospitalization time, type of services, required QoS, etc. Figure 6 shows the patient journey in the NICU at SickKids hospital. SickKids has 36 NICU beds including different types of patients. Depending on the type of patients, different numbers of algorithms for various periods will be triggered.

After discharging of a patient, a new patient will be submitted to NICU in 4-6 hours. Fifty percent of patients are term babies who are referred to SickKids for surgical purposes. Surgical babies stay in hospital for 5 days approximately, and 8 medical algorithms will be applied for after-surgery monitoring. The rest of patients, that is, preterm babies, are classified into three categories: babies who are born at 32-35, 27-32, and 23-27 weeks of their gestation age. The first group (ie, 30% of the patients) will be monitored by at most 8 medical algorithms for a mean period of 8 days. The second group (15%) of preterm babies will be monitored by 10 or fewer algorithms for an average time of 1 month. The third group is divided into two subclasses depending on medical conditions: 80% of this group (ie, 4% of the whole population) needs to be monitored by 20 or more algorithms for 4 months, and 20% (ie, 1% of the whole population) needs to be monitored by 20 or more algorithms for approximately 6 months. As Figure 6 suggests, SickKids NICU can be modeled as a single heterogeneous finite queue with multiple service facilities. Each type of patient has distinct characteristics in terms of length of stay and number of algorithms. Algorithms are also different in terms of required computational resources.

The SickKids NICU receives more admission requests than it has space for and prioritizes neonatal surgical patients. Other patients are typically redirected to either Sunnybrook Hospitals or Mount Sinai Hospital's NICU when SickKids is operating at or near capacity. The total number of bed spaces available for admission is thus 118, with 40 and 42 of these spaces available at these other 2 hospitals, respectively. We model the Artemis-IC platform as an M/G/m/m queuing system (*M* stands for Markovian, ie, Poisson), which indicates that the interarrival time of patient's arrival is exponentially distributed with the mean value of λ while patients' resident time at NICU is independently and identically distributed random variables that follows a general distribution. The system under consideration contains *m* servers (ie, bed spaces) that renders service in order of patients' arrivals (first-in-first-serve [FIFS]). The capacity of system is *m*, which means there is no extra room for queuing patients. As the population size of newborns is relatively high while the probability that a given newborn baby to be preterm is relatively small, the arrival process can be modeled as a

Poisson process. The details of the performance modeling can be found in [24].

Figure 6. Types of patients and their medical service path at SickKids NICU.



Results

The analytical model has been implemented in Maple 17 [29] in order to obtain the numerical results. First, we characterize the performance metrics for the current configuration of Artemis-IC at SickKids that was described earlier in the section. Table 1 shows the performance metrics and important exogenous parameters. The average length of stay for patients is 16 days, and each patient requires 9 algorithms on average on the stream computing platform (ie, IBM Streams). The mean number of monitored patients (ie, occupancy rate) is 34.9, so that 311 algorithms will be running on Streams. Each algorithm is consuming approximately 110 MB of memory, which indicates the requirement of at least 32 GB of memory for the stream-computing cluster. Note that this amount of memory is just for application hosts and the management hosts require at least 2 GB more of memory.

Table 1. Configuration parameters and performance metrics for current capacity of SickKids NICU.

Parameter	Value
Beds in NICU, n	36
Patient arrival (patient/day), mean rate	4.5
Length of stay for patients (days), mean	16
Number of algorithms for 1 patient, mean	9
All running algorithms on Streams, n	311
NICU's service (patient/day), rate	0.062
Blocking probability	0.455
Number of patients in NICU, mean	34.9
Memory per algorithm, mean MB	110
Required memory on Streams cluster, GB	32
Required CPU cores for Streams cluster, n	16
Required storage for a patient's data (per day), MB	700
Required storage on BigInsights cluster (per year), TB	8.6

As can be seen in Table 1, the amount of minimum storage for the Hadoop cluster (ie, BigInsights cluster) to only support the accommodation of raw physiological data for 1 year is 8.6 TB. Depending on the data schema design on the BigInsights cluster, additional storage might be required for the metadata. Moreover, the storage required for nonphysiological data such as patient information, laboratory results, and other related medical data should be added on top of this calculation.

Figure 7 shows the amount of storage for the BigInsights cluster, for 10, 36, 50, 60, 70-120 beds in the NICU. Note that this amount is only for raw physiological data acquired from NICU. The amount of storage increases linearly with respect to NICU capacity up to 60 beds. Then between 60 and 80 beds, it is increases sublinearly and in the end flattens. After reaching the capacity of 90 beds, the amount of required storage remains unchanged, which indicates that the NICU entered into the unsaturated regime and can accommodate all new patient arrivals. In other words, for 1 year, 16 TB of storage is sufficient for the SickKids NICU regardless of NICU's capacity (ie, the number of bed spaces).

We are also interested in studying the number of patients who get blocked, that is, redirected to another NICU, due to the capacity limitations of the NICU of interest. To this end, we characterize the blocking probability for the NICU with the capacity of 10-120 beds. As can be seen in Figure 8, for the current capacity of SickKids NICU (ie, 36 beds), 46% of patients get blocked. However, by increasing the capacity to 150 beds, the blocking will be less than 1%.

We also investigated the amount of memory and computation power for the stream-computing cluster for different configurations. Figure 9 shows the trend of required memory and number of CPU cores with respect to number of beds. For up to 70 beds, there is a linear dependency between the required memory and capacity; however, results show 60 GB of memory suffices for the Streams cluster based on these arrival and departure rates.

Our calculation for computation power is based on the standard CPU cores, that is, 2.00-GHz core, on IBM Softlayer cloud-based servers [30] and our experiments, which revealed that for each 20 algorithms we need a dedicated CPU core. The trend for computation power is almost similar to memory, explained above. We shall repeat the fact that these amounts of memory and computation power are just for application hosts. Depending on the deployment of management servers, extra resources might be needed.

Figure 7. Required storage for BigInsights cluster for different configurations.

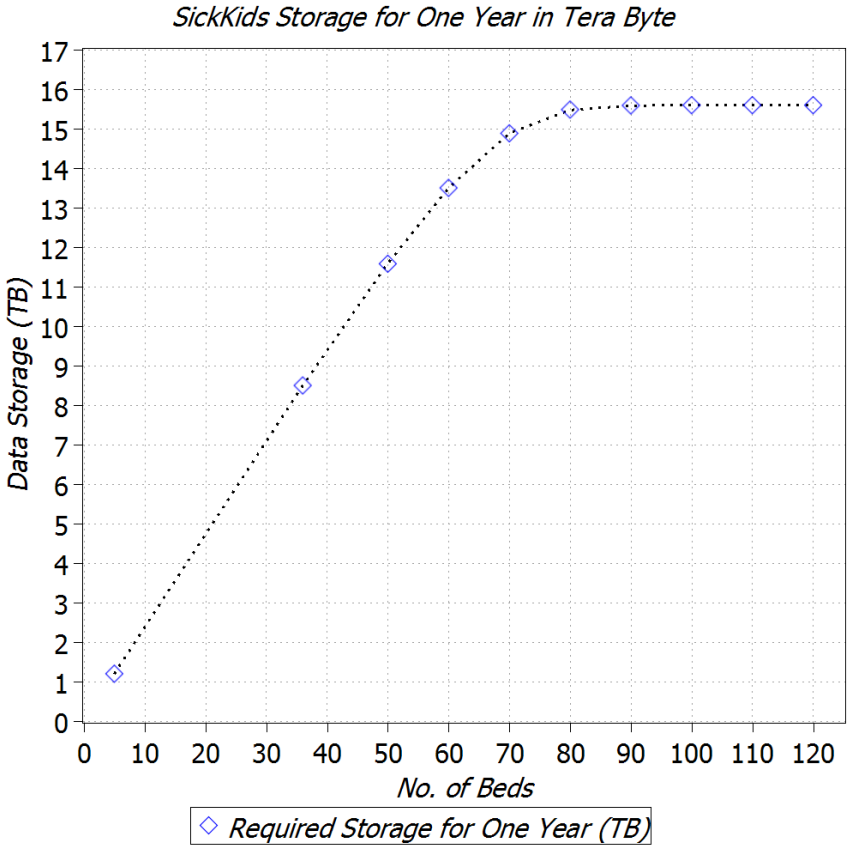


Figure 8. Blocking probability for different configurations.

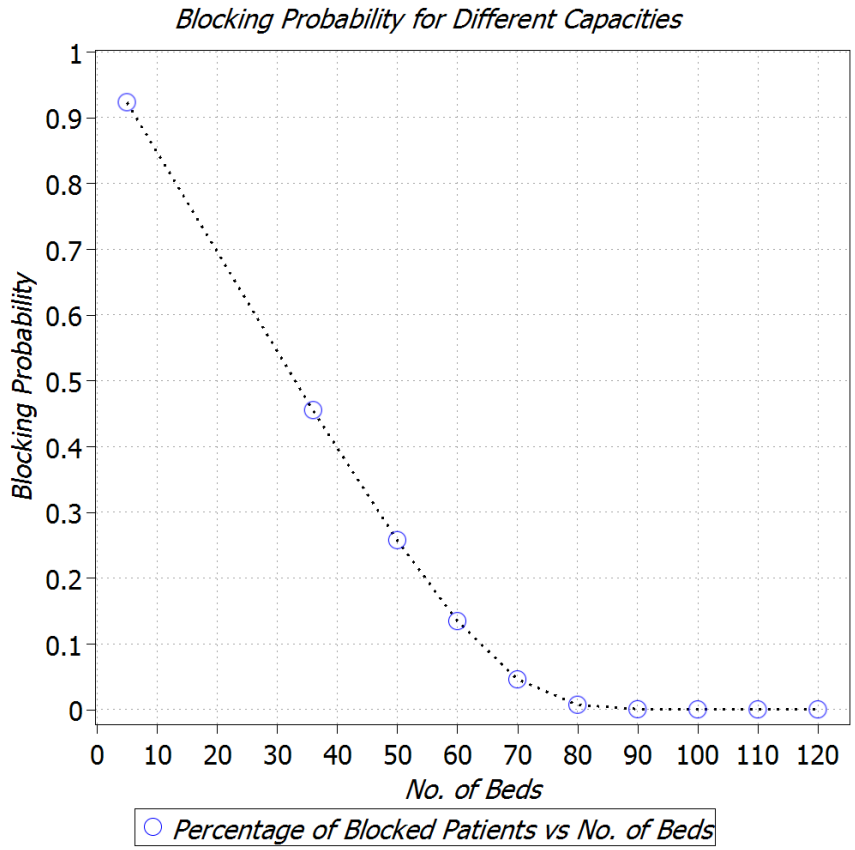
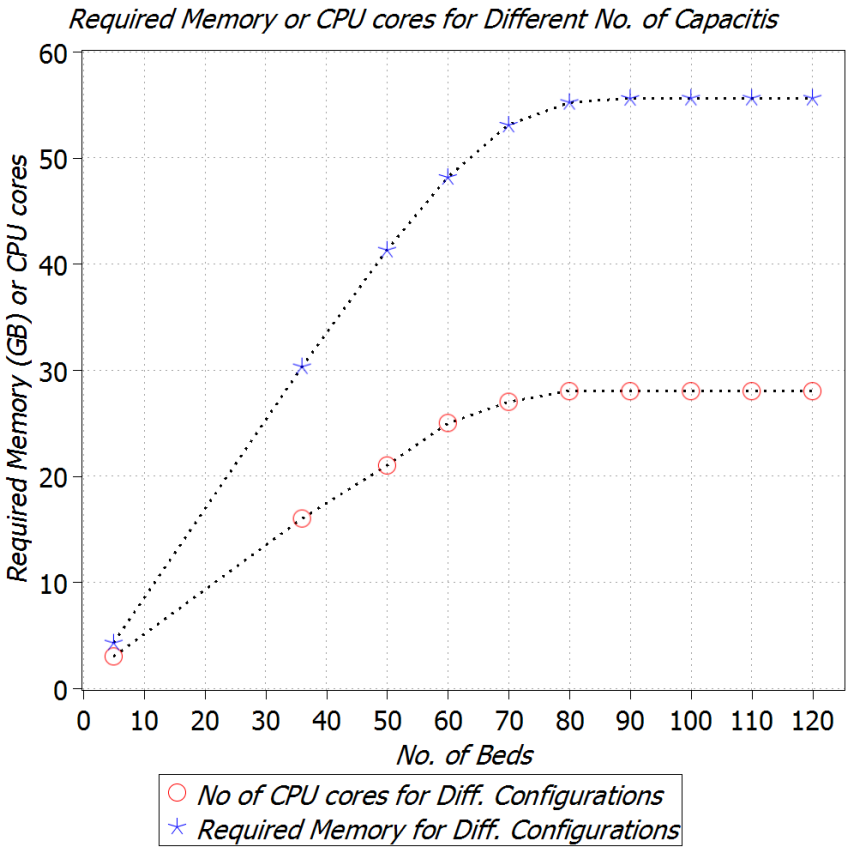


Figure 9. Required memory and computation power for Streams cluster for different configurations.



Discussion

Principal Considerations

We have described and evaluated the design, implementation, and pilot deployments of a framework that provides health analytics as services. This framework can be considered as a general architecture that can be tailored for different use cases in the health informatics domain. One such customization is the Artemis-IC project that provides a way for clinicians to have online, real-time execution of the clinical rules in an intensive care environment. Moreover, Artemis-IC provides researchers with a rich set of easy access data and analytics tools by which knowledge discovery will be much more attainable than in the past. Because Artemis-IC's target environments are critical care units, we have carried out extensive performance evaluation in order to guarantee expected quality of service and a high level of availability in particular. This work has three main aspects to be compared with similar works in the area, namely, data collection, real-time, and retrospective analysis. In the following sections, we compare our research to related work with regard to these three aspects.

Data Collection

Collection of the physiological data is the first step in the development of a CDSS. As technology has progressed, the amount of physiological data as well as clinical information about patients, for example, EHR, has grown significantly [31]. As such, developing systems that record these data securely and at a suitable sampling rate and make them highly available is a research topic on its own [26,32,33].

Sukuvaara et al [34] developed a system called *DataLog*, which would connect to bedside monitors through an RS232 serial interface to collect physiological signals every 5 seconds. They performed some trending analysis on the signals and combined it with heuristic "if-then" rules to create a knowledge-based alarm system. However, capturing a data point once every 5 seconds is not enough to implement complex algorithms in the real-time environment, which is a part of our solution. In addition, only numeric signals are collected with DataLog, and no waveform data are captured, which is an important component of detecting conditions in real time.

Moody et al [35] developed customized software to log the signals coming from the Hewlett Packard content management system (Merlin) bedside monitors that were being used in the medical, surgical, and cardiac ICUs of Beth Israel Hospital, Boston, using a pair of RS232 serial interface cards in the monitor and communicating the data to a standard personal computer over a serial interface. They were able to record 3 ECG signals each sampled at 500 Hz and 4 or 5 other signals sampled at 125 Hz, in addition to periodic measurements and alarm messages. While the amount of data collected is impressive, their approach was to strictly record and store the data for the purpose of retrospective analysis. There was no functionality to serve the data for any online processing.

Saeed et al [36] designed a system that collected physiological and clinical data from the information management system on the hospital's local area network for creating a temporal ICU patient database called *MIMIC II*. They monitored patients admitted to an 8-bed medical ICU and an 8-bed coronary care unit. The physiological data consisted of 4 continuously

monitored waveforms (2 leads of ECG, arterial blood pressure, pulmonary artery pressure) sampled at 125 Hz, 1-minute parameters (heart rate, blood pressure, oxygen saturation, and cardiac output), as well as monitor-generated alarms. The strength in their approach is the ability to vary the presentation of data depending on the specific type of research for which the data are being used. Users of the database can extract a detailed record of a single signal, or more temporal analysis data from many signals can be displayed in one view. However, this ability to provide data temporally can be done only after considerable preprocessing and data fusion and is inherently retrospective.

A pilot and customized implementation of our method (ie, Artemis-IC) in SickKids Hospital is capable of collecting 15 data streams including 12 scalars (reading 1 integer per second) and 3 waveform streams (reading 60 doubles per second) and ECG (reading 512 double per second). In addition, the Artemis-IC clinical information system (CIS) adapter interfaces with the clinical information management system (CIMS) to access the SickKids CIMS patient EHR and stream the data into the framework [22].

Real-Time Patient Monitoring

Current cutting-edge health informatics research projects aim to discover new condition onset behaviors that are evident in physiological data streams earlier than traditional detection of conditions in critical care data [23]. To this end, some hospitals may participate in pilot programs that aim to collect real-time patient data from network-enabled monitoring devices. These collected data are then analyzed to extract relevant temporal behaviors and usually stored for future data mining and analysis operations.

Historically, physiological stream monitoring of ICU patients has been provided by “black box” regulatory body-approved medical devices located at the patients’ bedside. While there has been a growing body of biomedical engineering and clinical research over the past 20-30 years proposing newer approaches for advanced physiological stream monitoring, they still predominantly have a physiological stream, clinical condition, or patient-centric approach [37]. Zhang et al [38] have discussed the implementation of a Health Data Stream Analytics System called the “Anesthetics Data Analyzer,” which has been developed to provide anesthetists with the ability to monitor and query trends in physiological signals data, a kind of stream data from the health care domain.

The BioStream [39] research project was designed to support the continuous monitoring of heart information of a patient on top of a general-purpose stream processing software architecture. The ECG was the main signal of interest. The goal of the group was to develop the prototype and collaborate with a medical institution on a pilot study. A Drexel University research team set out to design a system that performed online continuous processing of an ICU patient’s data stream and data capture to perform offline analysis to develop new clinical hypotheses [40].

As we propose a programmable component for the real-time processing in our solution, it can be customized to track a vast variety of diseases simultaneously. This capability is in part because of a comprehensive data collection followed by efficient ETL techniques that we employed in the design and implementation process. Moreover, there exist five active studies for developing and certifying medical algorithms to be deployed on the real-time component.

Retrospective Analysis and Knowledge Discovery

The taxonomy for analytic workflow systems has already been presented [41]. Based on the taxonomy and a study of the existing analytic software and systems, the authors proposed the conceptual architecture of CLoud-based Analytics-as-a-Service (CLAaaS). They outline the features that are important for CLAaaS as a service provisioning system such as user- and domain-specific customization and assistance, collaboration, modular architecture for scalable deployment, and service level agreement (SLA). We considered the aforementioned outlined features for designing the proposed framework in this work.

Analytics have been utilized in various aspects of health care including predictive risk assessment, clinical decision support, home health monitoring, finance, and resource allocation [6]. The proliferation of big data and analytics in health care has spawned a growing demand for clinical informatics professionals who can bridge the gap between the medical and information sciences.

John Tukey pioneered the use of exploratory data analysis nearly four decades ago [42]. Various packages and languages that support exploratory data analysis have been developed since. This includes S, S-Plus, R, SPSS, SAS, OLAP, and MATLAB [43,44]. A recent view of modern data exploration practices is available from Behrens and Yu [45]. All these approaches can be used as the knowledge discovery engine in our proposed architecture.

The retrospective analysis of previously persistently stored physiological data through the determination and assessment of TA-based qualitative behaviors from the analysis of quantitative physiological data has been widely employed. However, research is either physiological stream-clinical condition or patient centric [1]. A structured approach for the translation of the knowledge gained from this research, which is predominantly statistical and sometimes more recently data mining in nature, has been lacking [37,46].

One approach to the Software-as-a-Service utilizes the SOA approach to software design where software services are made available to the cloud through a series of Web services. Examples of early work showing the potential for the use of cloud computing in health care are emerging [11,47]; however, these research efforts do not provide functional support to critical care. McGregor [48,49] proposes a functional set of Web services to support critical care as part of her solution manager service as applied to health care. However, aspects such as rule definition are not clearly defined within that functional set. The application of cloud computing for the provision of a service of critical care supporting both real-time patient monitoring and retrospective clinical research remains an open research problem.

Strengths and Limitations

One of the main strengths of our work is the openness of the proposed framework. It is general enough to be applied to various scenarios in health informatics. The stream computing platform in the clinical edition can be programmed for monitoring different types of patients including but not limited to neonates, children, adults, and the elderly in critical care units, home, work, and even in transit. Any medical diagnostic approach that can be described algorithmically can be deployed and programmed on a real-time processing unit. Another key strength of the framework is the modular design of the architecture. In the research edition, any interested big data solution can be utilized. For example, any Hadoop distribution (eg, Cloudera [50], Hortonworks [51]) or other big data analytics tools such as Spark [52] can be employed for different types of retrospective analytics, provided that different types of analytics such as machine learning, statistical modeling, batch processing, interactive, streaming, graph, and in-memory analysis are accessible to researchers. In addition, our experience in customization of the framework for the NICU revealed that it could be deployed with minimum intervention with current procedures and policies. For example, for Artemis-IC deployment at SickKids we used only the spare port at the bedside monitors. We also developed an interface to interact with the clinical management information system to get the EHR from the hospital. Moreover, the systematic performance modeling can be easily extended or customized to support other medical care units. Estimation and prediction of the appropriate underlying infrastructure is no longer an unknown question.

However, there exist some limitations that need to be addressed properly and according to the target deployment. First and foremost is adopting appropriate privacy mechanisms for the physiological and medical data. For Artemis-IC, we used a simple deidentification technique that might not be completely secure and efficient. We use this technique to enable a simple reidentification process at hospitals. A more robust approach may apply encryption and perform analytics on encrypted data [53]. A second challenge is the ETL process for physiological data. This process should eliminate noise inputs from valid data efficiently; this is a research topic on its own [54,55]. Third, the process of medical algorithms certification is a complex and time-consuming process that prevents acquiring actual benefits out of the system in a timely manner. In other words, the lack of standardization seems to be an obstacle toward the adoption of systems such as Artemis-IC.

Conclusion

Our work fills the gap by providing a solution that can utilize the latest achievements in cloud-based analytics for health care informatics; it provides both real-time and retrospective analysis capabilities for various stakeholders. Moreover, we proposed a performance model that can be used for the capacity planning of the Artemis-IC in advance of its physical deployment. Artemis-IC and the corresponding performance model can be tailored for other ICUs as well; the architecture is plug-in–based so that similar open-source or commercial components can be integrated to realize the solution. Artemis-IC can also be deployed on any other cloud environment (ie, cloud agnostic).

Acknowledgments

This research is supported by Southern Ontario Smart Computing Innovation Platform (SOSCIP), the Canada Research Chairs Program (No 950-225945), and the Canadian Foundation for Innovation (No 203427).

Conflicts of Interest

None declared.

References

1. Tsoukalas A, Albertson T, Tagkopoulos I. From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis. JMIR Med Inform 2015;3(1):e11 [FREE Full text] [doi: 10.2196/medinform.3445] [Medline: 25710907]
2. Gunter TN, Terry NP. The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. J Med Internet Res 2005;7(1):e3 [FREE Full text] [doi: 10.2196/jmir.7.1.e3] [Medline: 15829475]
3. Fonkych KR, Taylor RC. The state and pattern of health information technology adoption. Santa Monica, CA: Rand Corp; 2005.
4. Kellermann AS. What it will take to achieve the as-yet-unfulfilled promises of health information technology. Health Aff 2013;32(1):63-68. [doi: 10.1377/hlthaff.2012.0693]
5. Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: a new business paradigm for biomedical information sharing. J Biomed Inform 2010 Apr;43(2):342-353 [FREE Full text] [doi: 10.1016/j.jbi.2009.08.014] [Medline: 19715773]
6. Simpao A, Ahumada L, Gálvez JA, Rehman M. A review of analytics and clinical informatics in health care. J Med Syst 2014 Apr;38(4):45. [doi: 10.1007/s10916-014-0045-x] [Medline: 24696396]
7. Raghupathi WV. An overview of health analytics. J Health Med Informat 2013;4(132):2. [doi: 10.4172/2157-7420]
8. Chawla ND, Davis DA. Bringing big data to personalized healthcare: a patient-centered framework. J Gen Intern Med 2013 Sep;28 Suppl 3:S660-S665 [FREE Full text] [doi: 10.1007/s11606-013-2455-8] [Medline: 23797912]

9. Miriovsky BJ, Shulman LN, Abernethy AP. Importance of health information technology, electronic health records, and continuously aggregating data to comparative effectiveness research and learning health care. J Clin Oncol 2012 Dec 1;30(34):4243-4248. [doi: 10.1200/JCO.2012.42.8011] [Medline: 23071233]
10. Wharam JJ, Weiner JP. The promise and peril of healthcare forecasting. Am J Manag Care 2012 Mar;18(3):e82-e85 [FREE Full text] [Medline: 22435964]
11. Bhavani R, Senthilkumar J, Chilambuchelvan G, Manjula D, Krishnamoorthy R, Kannan A. Cimidx: Prototype for a cloud-based system to support intelligent medical image diagnosis with efficiency. JMIR Med Inform 2015 Mar;3(1). [doi: 10.2196/medinform.3709]
12. Rojas CC, Patton RM, Beckerman BG. Characterizing mammography reports for health analytics. J Med Syst 2011 Oct;35(5):1197-1210. [doi: 10.1007/s10916-011-9685-2] [Medline: 21671070]
13. McGregor C. System, method and computer program for multi-dimensional temporal data mining. US Patent 8583686 2013 Nov.
14. Chan M, Estève D, Fourniols J, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
15. Banaee H, Ahmed MU, Loutfi A. Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. Sensors (Basel) 2013;13(12):17472-17500 [FREE Full text] [doi: 10.3390/s131217472] [Medline: 24351646]
16. Tseng KC, Hsu CL, Chuang YH. Designing an intelligent health monitoring system and exploring user acceptance for the elderly. J Med Syst 2013 Dec;37(6):9967. [doi: 10.1007/s10916-013-9967-y] [Medline: 24037138]
17. Baig MH, Gholamhosseini H. Smart health monitoring systems: an overview of design and modeling. J Med Syst 2013 Apr;37(2):9898. [doi: 10.1007/s10916-012-9898-z] [Medline: 23321968]
18. Delen DH, Demirkan H. Data, information and analytics as services. Decis Support Syst 2013 Apr;55(1):359-363. [doi: 10.1016/j.dss.2012.05.044]
19. Catley C, Smith K, McGregor C, James A, Eklund JM. A framework to modeltranslate clinical rules to support complex real-time analysis of physiologicalclinical data. 2010 Presented at: 1st ACM International Health Informatics Symposium; 2010; Washington, DC p. 307-315. [doi: 10.1145/1882992.1883036]
20. Kramer M, Platt R, Yang H, Joseph K, Wen S, Morin L, et al. Secular trends in preterm birth: a hospital-based cohort study. JAMA 1998 Dec 2;280(21):1849-1854. [Medline: 9846780]
21. McGregor C. Big data in neonatal intensive care. Computer 2013 Jun;46(6):54-59. [doi: 10.1109/MC.2013.157]
22. Blount M, Ebling MR, Eklund JM, James AG, McGregor C, Percival N, et al. Real-time analysis for intensive care: development and deployment of the artemis analytic system. IEEE Eng Med Biol Mag 2010;29(2):110-118. [doi: 10.1109/MEMB.2010.936454] [Medline: 20659848]
23. McGregor C. A cloud computing framework for real-time rural and remote service of critical care. 2011 Presented at: 24th International Symposium on Computer-Based Medical Systems (CBMS); Jun 27-30, 2011; Bristol, United Kingdom p. 1-6.
24. Khazaei H, McGregor C, Eklund M, El-Khatib K, Thommandram A. Toward a big data healthcare analytics system: a mathematical modeling perspective. 2011 Jun Presented at: IEEE 10th World Congress on Services (DSS); Jun 2014; Anchorage, Alaska.
25. McGregor C, Catley C, Padbury J, James A. Late onset neonatal sepsis detection in newborn infants via multiple physiological streams. J Crit Care 2013 Feb;28(1):e11-e12. [doi: 10.1016/j.jcrc.2012.10.037]
26. Thommandram A, Eklund J, McGregor C, Pugh J, James A. A rule-based temporal analysis method for online health analytics and its application for real-time detection of neonatal spells. 2014 Presented at: IEEE International Congress on Big Data (BigData Congress); Jun 2014; Anchorage, AL p. 470-477.
27. Cirelli J, McGregor C, Graydon B, James A. Analysis of continuous oxygen saturation data for accurate representation of retinal exposure to oxygen in the preterm infant. Stud Health Technol Inform 2013;183:126-131. [Medline: 23388268]
28. Naik T, Bressan N, James A, McGregor C. Design of temporal analysis for a novel premature infant pain profile using artemis. J Crit Care 2013 Feb;28(1):e4. [doi: 10.1016/j.jcrc.2012.10.024]
29. Maplesoft Inc. Maple 18. 2014 Mar. URL: http://www.maplesoft.com/ [accessed 2015-11-11] [WebCite Cache ID 6cyCcFZeZ]
30. Softlayer, an IBM company. 2015 Feb. URL: http://www.softlayer.com/ [accessed 2015-11-11] [WebCite Cache ID 6cyFI4Hpn]
31. Williams H, Spencer K, Sanders C, Lund D, Whitley A, Kaye J, et al. Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. JMIR Med Inform 2015;3(1):e3 [FREE Full text] [doi: 10.2196/medinform.3525] [Medline: 25586934]
32. Korosec L, Balenko K, Hagens S. Impact of information technology on information gaps in canadian ambulatory care encounters. JMIR Med Inform 2015;3(1):e1 [FREE Full text] [doi: 10.2196/medinform.4066] [Medline: 25595130]
33. Rocha P, Queirós A, Augusto F, Rodríguez YL, Cardoso C, Grade M, et al. Information persistence services designed to support home care. JMIR Med Inform 2015;3(1):e15 [FREE Full text] [doi: 10.2196/medinform.3699] [Medline: 25757462]
34. Sukuvaara T, Koski EM, Mäkivirta A, Kari A. A knowledge-based alarm system for monitoring cardiac operated patients--technical construction and evaluation. Int J Clin Monit Comput 1993 May;10(2):117-126. [Medline: 8366312]

35. Freudenthal FA, Van-Stuijvenberg M, Va-Goudoever J. A quiet NICU for improved infants health, development and well-being: a systems approach to reducing noise and auditory alarms. Cogn Technol Work 2013;15(3):329-345.
36. Saeed M, Lieu C, Raber G, Mark R. MIMIC II: a massive temporal ICU patient database to support research in intelligent patient monitoring. Comput Cardiol 2002;29:641-644. [Medline: 14686455]
37. Stacey MC, McGregor C. Temporal abstraction in intelligent clinical data analysis: a survey. Artif Intell Med 2007 Jan;39(1):1-24. [doi: 10.1016/j.artmed.2006.08.002] [Medline: 17011175]
38. Zhang Q, Pang C, Mcbride S, Hansen D, Cheung C, Steyn M. Towards health data stream analytics. 2010 Presented at: IEEE/ICME International Conference on Complex Medical Engineering (CME); 2010; Toronto, Canada p. 282-287.
39. Bar-Or A, Healey J, Kontothanassis L, Van Thong J. Biostream: A system architecture for real-time processing of physiological signals. In: IEMBS'04. 2004 Presented at: 26th Annual International Conference of the Engineering in Medicine and Biology Society; 2004; San Francisco, CA p. 3101-3104.
40. Han H, Ryoo H, Patrick H. An infrastructure of stream data mining, fusion and management for monitored patients. 2006 Presented at: 19th IEEE International Symposium on Computer-Based Medical Systems, CBMS; 2006; Maribor, Slovenia p. 461-468.
41. Zulkernine F, Martin P, Zou Y, Bauer M, Gwadry-Sridhar F, Aboulnaga A. Towards cloud-based analytics-as-a-service (claaas) for big data analytics in the cloud. 2013 Presented at: IEEE International Congress on Big Data (BigData Congress); 2013; Santa Clara, CA p. 62-69.
42. Tukey J. Exploratory data analysis. Reading, Mass: Addison-Wesley Pub. Co; 1977.
43. Chambers JT, Hastie T. Statistical models in S. New York: Chapman & Hall; 1992.
44. Salley C. Providing OLAP (on-line analytical processing) to user analysts: An IT mandate. In: White Paper. Santa Clara, CA: Arbor Software Corporation; 1993.
45. Behrens JC. Exploratory data analysis. In: Handbook of Psychology. New York: Wiley & Sons; 2003.
46. Catley C, Stratti H, McGregor C. Multi-dimensional temporal abstraction and data mining of medical time series data: Trends and challenges. 2008 Presented at: 30th Annual International Conference of the Engineering in Medicine and Biology Society, EMBS; 2008; Vancouver, BC p. 4322-4325.
47. Hoang DL, Chen L. Mobile cloud for assistive healthcare (MOCASH). 2010 Presented at: Asia-Pacific Services Computing Conference (APSCC); 2010; Hangzhou, China p. 325-332.
48. McGregor C. e-baby web services to support local and remote neonatal intensive care. 2005 Presented at: HIC 2005, Thirteenth National Health Informatics Conference and HINZ 2005, Fourth Health Informatics Conference; 2005; Melbourne, Australia p. 344-349.
49. Hayes G, Khazaei H, El-Khatib K, McGregor C, Eklund M. Design and analytical model of a platform-as-a-service cloud for healthcare. J Internet Tech (JIT) 2014;16(1):139-150.
50. Cloudera Inc. 2015. URL: http://www.cloudera.com/content/cloudera [accessed 2015-11-11] [WebCite Cache ID 6cyEnZVJH]
51. Hortonworks Inc. 2015. URL: http://hortonworks.com/ [accessed 2015-11-11] [WebCite Cache ID 6cyEu3Nkb]
52. Apache Spark. 2015. URL: https://spark.apache.org/ [accessed 2015-11-11] [WebCite Cache ID 6cyF1S45h]
53. Bos J, Lauter K, Naehrig M. Private predictive analysis on encrypted medical data. J Biomed Inform 2014 Aug;50:234-243. [doi: 10.1016/j.jbi.2014.04.003] [Medline: 24835616]
54. Nizami S, Green J, McGregor C. Implementation of artifact detection in critical care: A methodological review. IEEE Rev Biomed Eng 2013;6:127-142.
55. Percival J, McGregor C, Percival N, James A. Enabling the integration of clinical event and physiological data for real-time and retrospective analysis. Lect Notes Bus Inf 2014:1-19.

Abbreviations

Artemis-IC: Artemis In Cloud
CDSS: clinical decision support system
CE: clinical edition
CIMS: Clinical Information Management System
CLAaaS: CLoud-based Analytic-as-a-Service
EHR: electronic health record
ETL: extract, transform, load
HAaaS: Health-Analytic-as-a-Service
ICU: intensive care unit
NICU: neonatal intensive care unit
QoS: quality of service
RE: research edition
SLA: service level agreement
SOA: service-oriented architecture
SPL: Stream Processing Language

SSL: secure socket layer
TA: temporal abstraction
UML: Unified Modeling Language
WIHRI: Women and Infants Hospital of Rhode Island

Edited by G Eysenbach; submitted 11.05.15; peer-reviewed by S Jagatheesan, I Kunnamo, R Vander Stichele; comments to author 29.07.15; revised version received 04.09.15; accepted 30.09.15; published 18.11.15.

Please cite as:
Khazaei H, McGregor C, Eklund JM, El-Khatib K
Real-Time and Retrospective Health-Analytics-as-a-Service: A Novel Framework
JMIR Med Inform 2015;3(4):e36
URL: http://medinform.jmir.org/2015/4/e36/
doi:10.2196/medinform.4640
PMID:26582268

Viewpoint

Disrupting Electronic Health Records Systems: The Next Generation

Leo Anthony Celi[1*], MD, MPH, MS; Jeffrey David Marshall[2], MD, MPH, MS; Yuan Lai[3], MSAUSI, MUP; David J Stone[4*], MD

[1]Beth Israel Deaconess Medical Center, Division of Pulmonary, Critical Care, and Sleep Medicine, Boston, MA, United States

[2]Yale University, Yale-New Haven Hospital, Department of Pulmonary and Critical Care, New Haven, CT, United States

[3]Center for Urban Science and Progress, New York University, New York, NY, United States

[4]Center for Wireless Health, Departments of Anesthesiology and Neurological Surgery, University of Virginia, Charlottesville, VA, United States

[*]these authors contributed equally

Corresponding Author:
Leo Anthony Celi, MD, MPH, MS
Beth Israel Deaconess Medical Center
Division of Pulmonary, Critical Care, and Sleep Medicine
330 Brookline Avenue
Boston, MA, 02215
United States
Phone: 1 617 667 5864
Fax: 1 617 667 1604
Email: lceli@bidmc.harvard.edu

Abstract

The health care system suffers from both inefficient and ineffective use of data. Data are suboptimally displayed to users, undernetworked, underutilized, and wasted. Errors, inefficiencies, and increased costs occur on the basis of unavailable data in a system that does not coordinate the exchange of information, or adequately support its use. Clinicians' schedules are stretched to the limit and yet the system in which they work exerts little effort to streamline and support carefully engineered care processes. Information for decision-making is difficult to access in the context of hurried real-time workflows. This paper explores and addresses these issues to formulate an improved design for clinical workflow, information exchange, and decision making based on the use of electronic health records.

(JMIR Med Inform 2015;3(4):e34) doi:10.2196/medinform.4192

KEYWORDS

clinical decision making; clinical decision support; electronic health records; electronic notes

Introduction

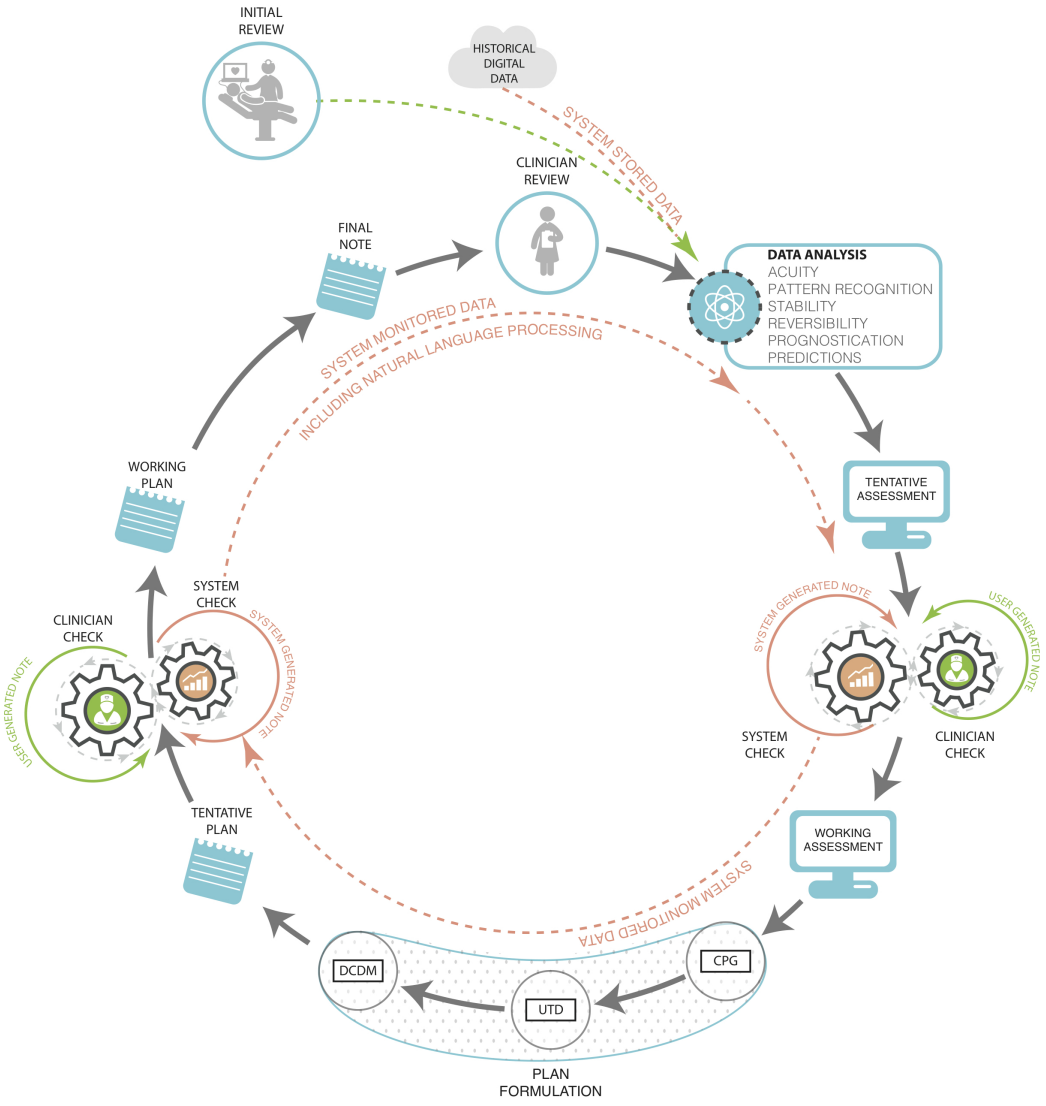
Weed introduced the "Subjective, Objective, Assessment, and Plan" (SOAP) note in the late 1960s [1]. This note entails a high-level structure that supports the thought process that goes into decision-making: subjective data followed by ostensibly more reliable objective data employed to formulate an assessment and subsequent plan. The flow of information has not fundamentally changed since that time, but the complexities of the information, possible assessments, and therapeutic options certainly have greatly expanded. Clinicians have not heretofore created anything like an optimal data system for medicine [2,3]. Such a system is essential to streamline workflow and support decision-making rather than adding to the time and frustration of documentation [4].

What this optimal data system offers is not a radical departure from the traditional thought processes that go into the production of a thoughtful and useful note. However, in the current early stage digitized medical system, it is still incumbent on the decision maker/note creator to capture the relevant priors, and to some extent, digitally scramble to collect all the necessary updates. The capture of these priors is a particular challenge in an era where care is more frequently turned over among different caregivers than ever before. Finally, based on a familiarity of the disease pathophysiology, the medical literature and evidence-based medicine (EBM) resources, the user is tasked with creating an optimal plan based on that assessment. In this so-called digital age, the amount of memorization, search, and assembly can be minimized and positively supported by a

well-engineered system purposefully designed to assist clinicians in note creation and, in the process, decision-making.

Since 2006, use of electronic health records (EHRs) by US physicians increased by over 160% with 78% of office-based physicians and 59% of hospitals having adopted an EHR by 2013 [5,6]. With implementation of federal incentive programs, a majority of EHRs were required to have some form of built-in clinical decision support tools by the end of 2012 with further requirements mandated as the *Affordable Care Act* (ACA) rolls out [7]. These requirements recognize the growing importance of standardization and systematization of clinical decision-making in the context of the rapidly changing, growing, and advancing field of medical knowledge. There are already EHRs and other technologies that exist, and some that are being implemented, that integrate clinical decision support into their functionality, but a more intelligent and supportive system can be designed that capitalizes on the note writing process itself. We should strive to optimize the note creation process as well as the contents of the note in order to best facilitate communication and care coordination. The following sections characterize the elements and functions of this decision support system (Figure 1).

Figure 1. Clinician documentation with fully integrated data systems support. Prior notes and data are input for the following note and decisions. Machine analyzes input and displays suggested diagnoses and problem list, and test and treatment recommendations based on various levels of evidence: CPG – clinical practice guidelines, UTD – Up to Date®, DCDM – Dynamic Clinical Data Mining.



Incorporating Data

Overwhelmingly, the most important characteristic of the electronic note is its potential for the creation and reception of what we term “bidirectional data streams” to inform both decision-making and research. By bidirectional data exchange, we mean that electronic notes have the potential to provide data streams to the entirety of the EHR database and vice versa. The data from the note can be recorded, stored, accessed, retrieved, and mined for a variety of real-time and future uses. This process should be an automatic and intrinsic property of clinical information systems. The incoming data stream is currently produced by the data that is slated for import into the note according to the software requirements of the application and the locally available interfaces [8]. The provision of information from the note to the system has both short- and long-term

benefits: in the short term, this information provides essential elements for functions such as benchmarking and quality reporting; and in the long term, the information provides the afferent arm of the learning system that will identify individualized best practices that can be applied to individual patients in future formulations of plans.

Current patient data should include all the electronically interfaced elements that are available and pertinent. In addition to the usual elements that may be imported into notes (eg, laboratory results and current medications), the data should include the immediate prior diagnoses and treatment items, so far as available (especially an issue for the first note in a care sequence such as in the ICU), the active problem list, as well as other updates such as imaging, other kinds of testing, and consultant input. Patient input data should be included after verification (eg, updated reviews of systems, allergies, actual medications being taken, past medical history, family history, substance use, social/travel history, and medical diary that may include data from medical devices). These data priors provide a starting point that is particularly critical for those note writers who are not especially (or at all) familiar with the patient. They represent historical (and yet dynamic) *evidence* intended to inform decision-making rather than "text" to be thoughtlessly carried forward or copied and pasted into the current note.

Although the amount and types of data collected are extremely important, how it is used and displayed are paramount. Many historical elements of note writing are inexcusably costly in terms of clinician time and effort when viewed at a level throughout the entire health care system. Redundant items such as laboratory results and copy-and-pasted nursing flow sheet data introduce a variety of “chartjunk” that clutters documentation and makes the identification of truly important information more difficult and potentially even introduces errors that are then propagated throughout the chart [9,10]. Electronic systems are poised to automatically capture the salient components of care so far as these values are interfaced into the system and can even generate an active problem list for the providers. With significant amounts of free text and “unstructured data” being entered, EHRs will need to incorporate more sophisticated processes such as natural language processing and machine learning to provide accurate interpretation of text entered by a variety of different users, from different sources, and in different formats, and then translated into structured data that can be analyzed by the system.

Optimally, a fully functional EHR would be able to provide useful predictive data analytics including the identification of patterns that characterize a patient’s normal physiologic state (thereby enabling detection of significant change from that state), as well as mapping of the predicted clinical trajectory, such as prognosis of patients with sepsis under a number of different clinical scenarios, and with the ability to suggest potential interventions to improve morbidity or mortality [11]. Genomic and other “-omic” information will eventually be useful in categorizing certain findings on the basis of individual susceptibilities to various clinical problems such as sepsis, auto-immune disease, and cancer, and in individualizing diagnostic and treatment recommendations. In addition, an embedded data analytic function will be able to recognize a constellation of relatively subtle changes that are difficult or impossible to detect, especially in the presence of chronic co-morbidities (eg, changes consistent with pulmonary embolism, which can be a subtle and difficult diagnosis in the presence of long standing heart and/or lung disease) [12,13].

The data presentation section must be thoughtfully displayed so that the user is not overwhelmed, but is still aware of what elements are available, and directed to those aspects that are most important. The user then has the tools at hand to construct the truly cognitive sections of the note: the assessment and plan. Data should be displayed in a fashion that efficiently and effectively provides a maximally informationally rich and minimally distracting graphic display. The fundamental principle should result in a thoughtfully planned data display created on the ethos of “just enough and no more,” as well as the incorporation of clinical elements such as severity, acuity, stability, and reversibility. In addition to the now classic teachings of Edward Tufte in this regard, a number of new data artists have entered the field [14]. There is room for much innovation and improvement in this area, as medicine transitions from paper to a digital format that provides enormous potential and capability for new types of displays.

Integrating the Monitors

Bedside and telemetry monitoring systems have become an element of the clinical information system but they do not yet interact with the EHR in a bidirectional fashion to provide decision support. In addition to the raw data elements, the monitors can provide data analytics that could support real-time clinical assessment as well as material for predictive purposes apart from the traditional noisy alarms [15,16]. It may be less apparent how the reverse stream (EHR to bedside monitor) would work, but the EHR can set the context for the interpretation of raw physiologic signals based on previously digitally captured vital signs, patient co-morbidities and current medications, as well as the acute clinical context.

In addition, the display could provide an indication of whether technically ”out of normal range” vital signs (or labs in the emergency screen described below) are actually “abnormal” for this particular patient. For example, a particular type of laboratory value for a patient may have been chronically out of normal range and not represent a change requiring acute investigation and/or treatment. This might be accomplished by displaying these types of ”normally abnormal” values in purple or green rather than red font for abnormal, or via some other designating graphic. The purple font (or whatever display mode was utilized) would designate the value as technically abnormal, but perhaps not *contextually* abnormal. Such designations are particularly important for caregivers who are not familiar with the patient.

It also might be desirable to use a combination of accumulated historical data from the monitor and the EHR to formulate personalized alarm limits for each patient. Such personalized alarm limits would provide a smarter range of acceptable values for each patient and perhaps also act to reduce the unacceptable number of false positive alarms that currently plague bedside

caregivers (and patients) [17]. These alarm limits would be dynamically based on the input data and subject to reformulation as circumstances changed. We realize that any venture into alarm settings becomes a regulatory and potentially medico-legal issue, but these intimidating factors should not be allowed to grind potentially beneficial innovations to a halt. For example, "hard" limits could be built into the alarm machine so that the custom alarm limits could not fall outside certain designated values.

Supporting the Formulation of the Assessment

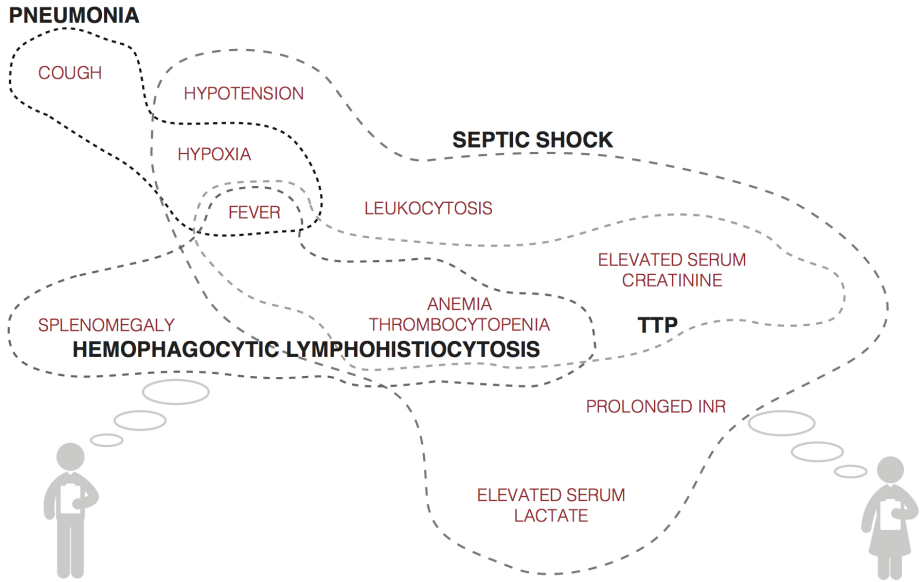
Building on both prior and new, interfaced and manually entered data as described above, the next framework element would consist of the formulation of the note in real time. This would consist of structured data so far as available and feasible, but is more likely to require real-time natural language processing performed on the free text being entered. Different takes on this kind of templated structure have already been introduced into several electronic systems. These include note templates created for specific purposes such as end-of-life discussions, or documentation of cardiopulmonary arrest. The very nature of these note types provides a robust context for the content. We also recognize that these shorter and more directed types of notes are not likely to require the kind of extensive clinical decision support (CDS) from which an admission or daily progress note may benefit.

Until the developers of EHRs find a way to fit structured data selection seamlessly and transparently into workflow, we will have to do the best we can with the free text that we have available. While this is a bit clunky in terms of data utilization purposes, perhaps it is not totally undesirable, as free text inserts a needed narrative element into the otherwise storyless EHR environment. Medical care can be described as an ongoing story and free text conveys this story in a much more effective and interesting fashion than do selected structured data bits. Furthermore, stories tend to be more distinctive than lists of structured data entries, which sometimes seem to vary remarkably little from patient to patient. But to extract the necessary information, the computer still needs a processed interpretation of that text. More complex systems are being developed and actively researched to act more analogously to our own "human" form of clinical problem solving [18], but until these systems are integrated into existing EHRs, clinicians may be able to help by being trained to minimize the potential confusion engendered by reducing completely unconstrained free text entries and/or utilizing some degree of standardization within the use of free text terminologies and contextual modifiers.

Employing the prior data (eg, diagnoses X, Y, Z from the previous note) and new data inputs (eg, laboratory results, imaging reports, and consultants' recommendations) in conjunction with the assessment being entered, the system would have the capability to check for inconsistencies and omissions based on analysis of both prior and new entries. For example, a patient in the ICU has increasing temperature and heart rate, and decreasing oxygen saturation. These continuous variables are referenced against other patient features and risk factors to suggest the possibility that the patient has developed a pulmonary embolism or an infectious ventilator-associated complication. The system then displays these possible diagnoses within the working assessment screen with hyperlinks to the patient's flow sheets and other data supporting the suggested problems (Figure 2). The formulation of the assessment is clearly not as potentially evidence-based as that of the plan; however, there should still be dynamic, automatic and rapid searches performed for pertinent supporting material in the formulation of the assessment. These would include the medical literature, including textbooks, online databases, and applications such as WebMD. The relevant literature that the system has identified, supporting the associations listed in the assessment and plan, can then be screened by the user for accuracy and pertinence to the specific clinical context. Another potentially useful CDS tool for assessment formulation is a modality we have termed dynamic clinical data mining (DCDM) [19]. DCDM draws upon the power of large sets of population health data to provide differential diagnoses associated with groupings or constellations of symptoms and findings. Similar to the process just described, the clinician would then have the ability to review and incorporate these suggestions or not.

An optional active search function would also be provided throughout the note creation process for additional flexibility—clinicians are already using search engines, but doing so sometimes in the absence of specific clinical search algorithms (eg, a generic search engine such Google). This may produce search results that are not always of the highest possible quality [20,21]. The EHR-embedded search engine would have its algorithm modified to meet the task as Google has done previously for its search engine [22]. The searchable TRIP database provides a search engine for high-quality clinical evidence, as do the search modalities within Up to Date, Dynamed, BMJ Clinical Evidence, and others [23,24].

Figure 2. Mock visualization of symptoms, signs, laboratory results, and other data input and systems suggestion for differential diagnoses.



Supporting the Formulation of the Plan

With the assessment formulated, the system would then formulate a proposed plan using EBM inputs and DCDM refinements for issues lying outside EBM knowledge. Decision support for plan formulation would include items such as randomized control trials (RCTs), observational studies, clinical practice guidelines (CPGs), local guidelines, and other relevant elements (eg, Cochrane reviews). The system would provide these supporting modalities in a hierarchical fashion using evidence of the highest quality first before proceeding down the chain to lower quality evidence. Notably, RCT data are not available for the majority of specific clinical questions, or it is not applicable because the results cannot be generalized to the patient at hand due to the study's inclusion and exclusion criteria [25]. Sufficiently reliable observational research data also may not be available, although we expect that the holes in the RCT literature will be increasingly filled by observational studies in the near future [16,26]. In the absence of pertinent evidence-based material, the system would include the functionality which we have termed DCDM, and our Stanford colleagues have termed the "green button" [19,27]. This still-theoretical process is described in detail in the references, but in brief, DCDM would utilize a search engine type of approach to examine a population database to identify similar patients on the basis of the information entered in the EHR. The prior treatments and outcomes of these historical patients would then be analyzed to present options for the care of the current patient that were, to a large degree, based on prior data. The efficacy of DCDM would depend on, among other factors, the availability of a sufficiently large population EHR database, or an open repository that would allow for the sharing of patient data between EHRs. This possibility is quickly becoming a reality with the advent of large, deidentified clinical databases such as that being created by the Patient Centered Outcomes Research Institute [26].

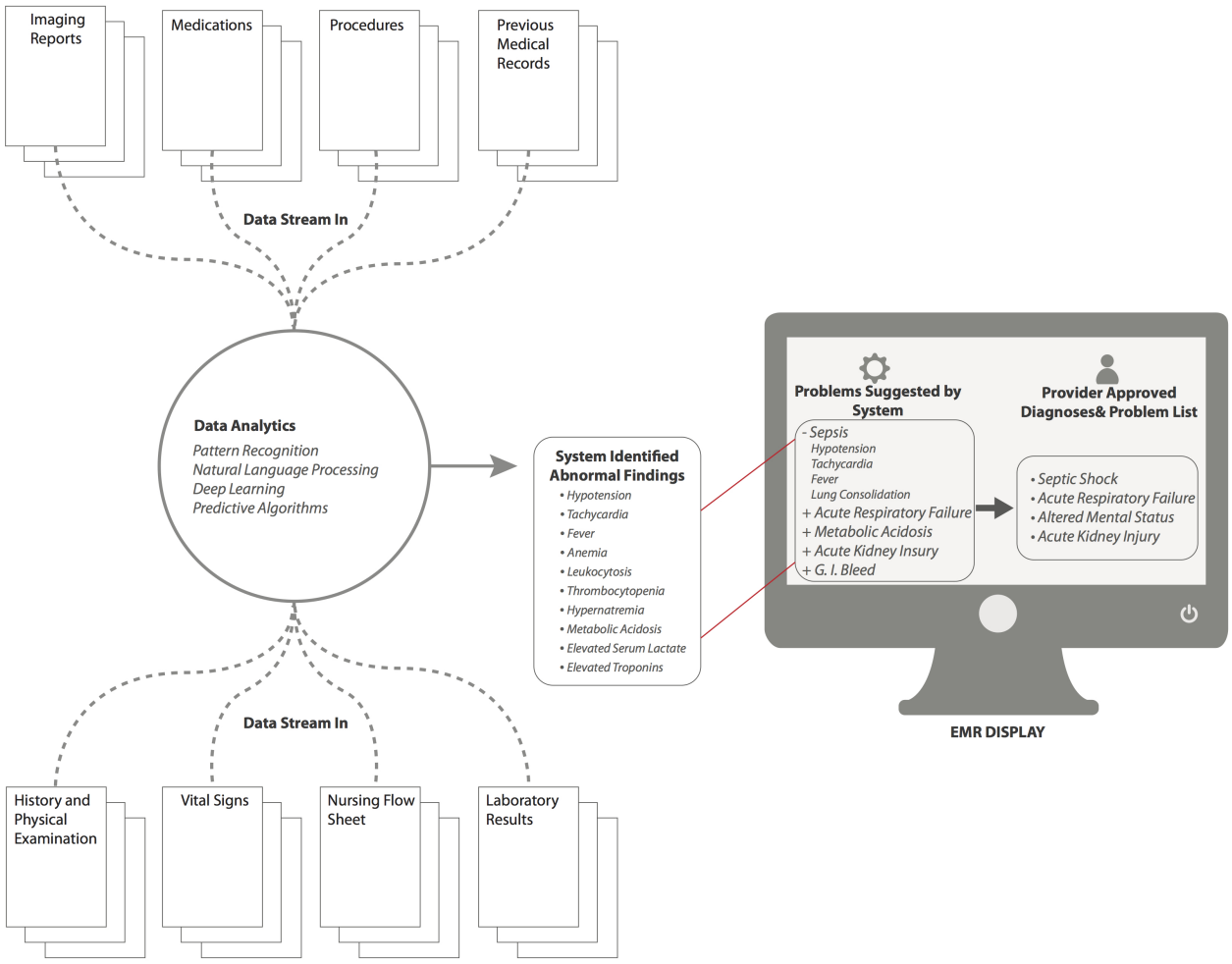
The tentative plan could then be modified by the user on the basis of her or his clinical "wetware" analysis. The electronic workflow could be designed in a number of ways that were modifiable per user choice/customization. For example, the user could first create the assessment and plan which would then be subject to comment and modification by the automatic system. This modification might include suggestions such as adding entirely new items, as well as the editing of entered items. In contrast, as described, the system could formulate an original assessment and plan that was subject to final editing by the user. In either case, the user would determine the final output, but the system would record both system and final user outputs for possible reporting purposes (eg, consistency with best practices). Another design approach might be to display the user entry *in toto* on the left half of a computer screen and a system-formulated assessment (Figure 3) and plan on the right side for comparison. Links would be provided throughout the system formulation so that the user could drill into EHR-provided suggestions for validation and further investigation and learning. In either type of workflow, the system would comparatively evaluate the final entered plan for consistency, completeness, and conformity with current best practices. The system could display the specific items that came under question and why. Users may proceed to adopt or not, with the option to justify their decision. Data reporting analytics could be formulated on the basis of compliance with EBM care. Such analytics should be done and interpreted with the knowledge that EBM itself is a moving target and many clinical situations do not lend themselves to resolution with the current tools supplied by EBM.

Since not all notes call for this kind of extensive decision support, the CDS material could be displayed in a separate columnar window adjacent to the main part of the screen where the note contents were displayed so that workflow is not affected. Another possibility would be an "opt-out" button by which the user would choose not to utilize these system

resources. This would be analogous but functionally opposite to the "green button" opt-*in* option suggested by Longhurst et al, and perhaps be designated the "orange button" to clearly make this distinction [27]. Later, the system would make a determination as to whether this lack of EBM utilization was justified, and provide a reminder if the care was determined to be outside the bounds of current best practices. While the goal is to keep the user on the EBM track as much as feasible, the system has to "realize" that real care will still extend outside those bounds for some time, and that some notes and decisions simply do not require such machine support.

There are clearly still many details to be worked out regarding the creation and use of a fully integrated *bidirectional* EHR. There currently are smaller systems that use some components of what we propose. For example, a large Boston hospital uses a program called QPID which culls all previously collected patient data and uses a Google-like search to identify specific details of relevant prior medical history which is then displayed in a user-friendly fashion to assist the clinician in making real-time decisions on admission [28]. Another organization, the American Society of Clinical Oncology, has developed a clinical Health IT tool called CancerLinQ which utilizes large clinical databases of cancer patients to trend current practices and compare the specific practices of individual providers with best practice guidelines [29]. Another hospital system is using many of the components discussed in a new, internally developed platform called Fluence that allows aggregation of patient information, and applies already known clinical practice guidelines to patients' problem lists to assist practitioners in making evidenced-based decisions [30]. All of these efforts reflect inadequacies in current EHRs and are important pieces in the process of selectively and wisely incorporating these technologies into EHRs, but doing so universally will be a much larger endeavor.

Figure 3. Mock screenshot for the "Assessment and Plan" screen with background data analytics. Based on background analytics that are being run by the system at all times, a series of "problems" are identified and suggested by the system, which are then displayed in the EMR in the box on the left. The clinician can then select problems that are suggested, or input new problems that are then displayed in the the box on the right of the EMR screen, and will now be apart of ongoing analytics for future assessment.



Conclusions

Medicine has finally entered an era in which clinical digitization implementations and data analytic systems are converging. We have begun to recognize the power of data in other domains and are beginning to apply it to the clinical space, applying digitization as a necessary but insufficient tool for this purpose (personal communication from Peter Szolovits, *The Unreasonable Effectiveness of Clinical Data. Challenges in Big Data for Data Mining, Machine Learning and Statistics Conference, March 2014*). The vast amount of information and

clinical choices demands that we provide better supports for making decisions and effectively documenting them. The Institute of Medicine demands a "learning health care system" where analysis of patient data is a key element in continuously improving clinical outcomes [31]. This is also an age of increasing medical complexity bound up in increasing financial and time constraints. The latter dictate that medical practice should become more standardized and evidence-based in order to optimize outcomes at the lowest cost. Current EHRs, mostly implemented over the past decade, are a first step in the digitization process, but do not support decision-making or streamline the workflow to the extent to which they are capable. In response, we propose a series of information system enhancements that we hope can be seized, improved upon, and incorporated into the next generation of EHRs.

There is already government support for these advances: The Office of the National Coordinator for Health IT recently outlined their 6-year and 10-year plans to improve EHR and health IT interoperability, so that large-scale realizations of this idea can and will exist. Within 10 years, they envision that we "should have an array of interoperable health IT products and services that allow the health care system to continuously learn and advance the goal of improved health care." In that, they envision an integrated system across EHRs that will improve not just individual health and population health, but also act as a nationwide repository for searchable and researchable outcomes data [32]. The first step to achieving that vision is by successfully implementing the ideas and the system outlined above into a more fully functional EHR that better supports both workflow and clinical decision-making. Further, these suggested changes would also contribute to making the note writing process an educational one, thereby justifying the very significant time and effort expended, and would begin to establish a true learning system of health care based on actual workflow practices. Finally, the goal is to keep clinicians firmly in charge of the decision loop in a "human-centered" system in which technology plays an *essential but secondary role.* As expressed in a recent article on the issue of automating systems [33]:

> *In this model (human centered automation)…technology takes over routine functions that a human operator has already mastered, issues alerts when unexpected situations arise, provides fresh information that expands the operator's perspective and counters the biases that often distort human thinking. The technology becomes the expert's partner, not the expert's replacement.*

Key Concepts and Terminology

A number of concepts and terms were introduced throughout this paper, and some clarification and elaboration of these follows:

- *Affordable Care Act (ACA)*: Legislation passed in 2010 that constitutes two separate laws including the *Patient Protection and Affordable Care Act* and the *Health Care and Education Reconciliation Act*. These two pieces of legislation act together for the expressed goal of expanding health care coverage to low-income Americans through expansion of Medicaid and other federal assistance programs [34].
- *Clinical Decision Support (CDS)* is defined by CMS as "a key functionality of health information technology" that encompasses a variety of tools including computerized alerts and reminders, clinical guidelines, condition-specific order sets, documentations templates, diagnostic support, and other tools that "when used effectively, increases quality of care, enhances health outcomes, helps to avoid errors and adverse events, improves efficiency, reduces costs, and boosts provider and patient satisfaction" [35].
- *Cognitive Computing* is defined as "the simulation of human thought processes in a computerize model…involving self learning systems that use data mining, pattern recognition and natural language processing to mimic the way the human brain works" [36]. Defined by IBM as computer systems that "are trained using artificial intelligence and machine learning algorithms to sense, predict, infer and, in some ways, think" [37].
- *Deep learning* is a form of machine learning (a more specific subgroup of *cognitive computing*) that utilizes multiple levels of data to make hierarchical connections and recognize more complex patterns to be able to infer higher level concepts from lower levels of input and previously inferred concepts [38]. Figure 3 demonstrates how this concept relates to patients illustrating the system recognizing patterns of signs and symptoms experienced by a patient, and then inferring a diagnosis (higher level concept) from those lower level inputs. The next level concept would be recognizing response to treatment for proposed diagnosis, and offering either alternative diagnoses, or change in therapy, with the system adapting as the patient's course progresses.
- *Dynamic clinical data mining (DCDM)*: First, data mining is defined as the "process of discovering patterns, automatically or semi-automatically, in large quantities of data" [39]. DCDM describes the process of mining and interpreting the data from large patient databases that contain prior and concurrent patient information including diagnoses, treatments, and outcomes so as to make real-time treatment decisions [19].
- *Natural Language Processing (NLP)* is a process based on machine learning, or deep learning, that enables computers to analyze and interpret unstructured human language input to recognize and even act upon meaningful patterns [39,40].

Conflicts of Interest

None declared.

References

1. Weed LL. Medical records, patient care, and medical education. Ir J Med Sci 1964 Jun;462:271-282. [Medline: 14160426]
2. Celi L, Csete M, Stone D. Optimal data systems: the future of clinical predictions and decision support. Curr Opin Crit Care 2014 Oct;20(5):573-580. [doi: 10.1097/MCC.0000000000000137] [Medline: 25137399]
3. Cook DA, Sorensen KJ, Hersh W, Berger RA, Wilkinson JM. Features of effective medical knowledge resources to support point of care learning: a focus group study. PLoS One 2013 Nov;8(11):e80318 [FREE Full text] [doi: 10.1371/journal.pone.0080318] [Medline: 24282535]
4. Cook DA, Sorensen KJ, Wilkinson JM, Berger RA. Barriers and decisions when answering clinical questions at the point of care: a grounded theory study. JAMA Intern Med 2013 Nov 25;173(21):1962-1969. [doi: 10.1001/jamainternmed.2013.10103] [Medline: 23979118]
5. Hsiao C, Hing E. US Department of Health and Human Services. Hyattsville, MD: US Department of Health and Human Services, CDC; 2001. Use and characteristics of electronic health record systems among office-based physician practices URL: http://www.cdc.gov/nchs/data/databriefs/db143.pdf [accessed 2014-12-30] [WebCite Cache ID 6VDPpho29]
6. US Department of Health and Human Services. 2014 Aug 14. More physicians and hospitals are using EHRs than ever before URL: http://www.hhs.gov/news/press/2014pres/08/20140807a.html [accessed 2014-12-30] [WebCite Cache ID 6VDPuIvPv]
7. HealthIT.gov. 2013 Apr 03. How to implement EHRs URL: http://www.healthit.gov/providers-professionals/ehr-implementation-steps [accessed 2014-12-30] [WebCite Cache ID 6VDPy8bEb]
8. Ghassemi M, Marshall J, Singh N, Stone DJ, Celi LA. Leveraging a critical care database: selective serotonin reuptake inhibitor use prior to ICU admission is associated with increased hospital mortality. Chest 2014 Apr;145(4):745-752 [FREE Full text] [doi: 10.1378/chest.13-1722] [Medline: 24371841]
9. Hammond KW, Helbig ST, Benson CC, Brathwaite-Sketoe BM. Are electronic medical records trustworthy? Observations on copying, pasting and duplication. AMIA Annu Symp Proc 2003:269-273 [FREE Full text] [Medline: 14728176]
10. Tufte E. The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press; 2001.
11. Tsoukalas A, Albertson T, Tagkopoulos I. From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis. JMIR Med Inform 2015;3(1):e11 [FREE Full text] [doi: 10.2196/medinform.3445] [Medline: 25710907]
12. Rochefort CM, Verma AD, Eguale T, Lee TC, Buckeridge DL. A novel method of adverse event detection can accurately identify venous thromboembolisms (VTEs) from narrative electronic health record data. J Am Med Inform Assoc 2015 Jan;22(1):155-165 [FREE Full text] [doi: 10.1136/amiajnl-2014-002768] [Medline: 25332356]
13. Amland RC, Dean BB, Yu H, Ryan H, Orsund T, Hackman JL, et al. Computerized clinical decision support to prevent venous thromboembolism among hospitalized patients: proximal outcomes from a multiyear quality improvement project. J Healthc Qual 2015;37(4):221-231. [doi: 10.1111/jhq.12069] [Medline: 26151096]
14. Papova M. Bloomberg Business Week. 2009. The 21 heroes of data visualization URL: http://images.businessweek.com/ss/09/08/0812_data_visualization_heroes/1.htm [accessed 2014-12-30] [WebCite Cache ID 6VDRAE5R9]
15. HeRO by MPSC. URL: http://www.heroscore.com/ [accessed 2014-12-30] [WebCite Cache ID 6VDRPJo4E]
16. Etiometry Inc. URL: http://www.etiometry.com/ [accessed 2014-12-30] [WebCite Cache ID 6VDRUwD7k]
17. Drew BJ, Harris P, Zègre-Hemsey JK, Mammone T, Schindler D, Salas-Boni R, et al. Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. PLoS One 2014;9(10):e110274 [FREE Full text] [doi: 10.1371/journal.pone.0110274] [Medline: 25338067]
18. Shen Y, Colloc J, Jacquet-Andrieu A, Lei K. Emerging medical informatics with case-based reasoning for aiding clinical decision in multi-agent system. J Biomed Inform 2015 Aug;56:307-317. [doi: 10.1016/j.jbi.2015.06.012] [Medline: 26133480]
19. Celi LA, Zimolzak AJ, Stone DJ. Dynamic clinical data mining: search engine-based decision support. JMIR Med Inform 2014;2(1):e13 [FREE Full text] [doi: 10.2196/medinform.3110] [Medline: 25600664]
20. Gold A. Fierce Health IT. 2014. Survey: Docs increasingly use online search engines to make clinical decisions, but Googling can also harm the patient-doctor relationship URL: http://www.fiercehealthit.com/story/survey-docs-increasingly-use-online-search-engines-make-clinical-decisions/2014-04-30 [accessed 2014-12-30] [WebCite Cache ID 6VDRc6H11]
21. Study: Google search results offer mixed-quality health information. 2014 Apr 14. iHealth Beat URL: http://www.ihealthbeat.org/articles/2014/4/14/study-google-search-results-offer-mixed-quality-health-information [accessed 2014-12-30] [WebCite Cache ID 6VDRhQtiq]
22. Sullivan D. Search Engine Land. FAQ: All about the new Google "Hummingbird" algorithm URL: http://searchengineland.com/google-hummingbird-172816 [accessed 2014-12-30] [WebCite Cache ID 6VDRmqWvR]
23. Trip Database. URL: https://www.tripdatabase.com/ [accessed 2015-09-10] [WebCite Cache ID 6bR7rpHJw]
24. Up To Date Clinical Database. URL: http://www.uptodate.com/home [accessed 2015-09-10] [WebCite Cache ID 6bR7uPaPb]
25. Humphreys K, Maisel NC, Blodgett JC, Fuh IL, Finney JW. Extent and reporting of patient nonenrollment in influential randomized clinical trials, 2002 to 2010. JAMA Intern Med 2013 Jun 10;173(11):1029-1031 [FREE Full text] [doi: 10.1001/jamainternmed.2013.496] [Medline: 23608926]

26. Selby JV, Lipstein SH. PCORI at 3 years - progress, lessons, and plans. N Engl J Med 2014 Feb 13;370(7):592-595. [doi: 10.1056/NEJMp1313061] [Medline: 24521104]
27. Longhurst CA, Harrington RA, Shah NH. A "green button" for using aggregate patient data at the point of care. Health Aff (Millwood) 2014 Jul;33(7):1229-1235. [doi: 10.1377/hlthaff.2014.0099] [Medline: 25006150]
28. QPID Health: Clinical Reasoning Power. URL: http://www.qpidhealth.com/ [accessed 2014-12-30] [WebCite Cache ID 6VDRrrOgg]
29. American Society of Clinical Oncology. Transforming cancer care through big data: ASCO unveils CancerLinQ prototype URL: http://www.asco.org/quality-guidelines/transforming-cancer-care-through-big-data-asco-unveils-cancerlinq-prototype [accessed 2014-12-30] [WebCite Cache ID 6VDRuCXpd]
30. UPMC Enterprises. Fluence: a breakthrough clinical platform for clinicians URL: http://upmctdc.com/Pages/fluence.aspx [accessed 2014-12-30] [WebCite Cache ID 6VDS3jptI]
31. Grossman C, Powers B, McGinnis M. The Institute of Medicine. Washington, DC: The National Academies Press; 2011 May 23. Digital infrastructure for the learning health system URL: http://www.iom.edu/Reports/2011/Digital-Infrastructure-for-a-Learning-Health-System.aspx [accessed 2014-12-30] [WebCite Cache ID 6VDSA0Mvl]
32. The Office of the National Coordinator for Health Information Technology. HealthIT.gov. 2014. Connecting health and care for the nation: a 10-year vision to achieve an interoperable health IT infrastructure URL: http://healthit.gov/sites/default/files/ONC10yearInteroperabilityConceptPaper.pdf [accessed 2014-12-30] [WebCite Cache ID 6VDSVW1Zn]
33. Carr N. Wall St Journal. 2014 Nov 21. Automation makes us dumb URL: http://www.wsj.com/articles/automation-makes-us-dumb-1416589342 [accessed 2014-12-30] [WebCite Cache ID 6VDSXtmh5]
34. Medicaid.gov. Affordable Care Act URL: http://medicaid.gov/affordablecareact/affordable-care-act.html [accessed 2015-03-24] [WebCite Cache ID 6XGguflhk]
35. CMS.gov - Centers for Medicare & Medicaid Services. Clinical decision support : more than just "alerts" tipsheet URL: http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/ClinicalDecisionSupport_Tipsheet-.pdf [accessed 2015-03-24] [WebCite Cache ID 6XGhTlwCC]
36. WhatIs.com. What is cognitive computing? URL: http://whatis.techtarget.com/definition/cognitive-computing [accessed 2015-03-24] [WebCite Cache ID 6XGh4btTm]
37. IBM Research. Cognitive computing URL: http://www.research.ibm.com/cognitive-computing/ [accessed 2015-03-24] [WebCite Cache ID 6XGh8lPLq]
38. Wikipedia. Deep learning URL: http://en.wikipedia.org/wiki/Deep_learning [accessed 2015-03-24] [WebCite Cache ID 6XGkO2Zi0]
39. Witten IH, Frank E, Hall MA. Data mining. In: Data Mining: Practical Machine Learning Tools and Techniques, Third Edition. Burlington, MA: Morgan Kaufmann; 2011.
40. Wikipedia. Natural language processing URL: http://en.wikipedia.org/wiki/Natural_language_processing [accessed 2015-03-24] [WebCite Cache ID 6XGkIOyfK]

Abbreviations

ACA: Affordable Care Act
CDS: clinical decision support
CPG: clinical practice guidelines
DCDM: dynamic clinical data mining
EBM: evidence-based medicine
EHR: electronic health record
RCT: randomized control trial
SOAP: subjective, objective, assessment, and plan

Edited by G Eysenbach; submitted 31.12.14; peer-reviewed by A Chang, C Rochefort; comments to author 23.02.15; revised version received 08.06.15; accepted 24.07.15; published 23.10.15.

Please cite as:
Celi LA, Marshall JD, Lai Y, Stone DJ
Disrupting Electronic Health Records Systems: The Next Generation
JMIR Med Inform 2015;3(4):e34
URL: http://medinform.jmir.org/2015/4/e34/
doi:10.2196/medinform.4192
PMID:26500106

Original Paper

Outcomes From Health Information Exchange: Systematic Review and Future Research Needs

William R Hersh[1], MD; Annette M Totten[1], PhD; Karen B Eden[1], PhD; Beth Devine[2], MBA, PhD, PharmD; Paul Gorman[1], MD; Steven Z Kassakian[1], MD; Susan S Woods[3], MD; Monica Daeges[1], BS; Miranda Pappas[1], MA; Marian S McDonagh[1], PharmD

[1]Pacific Northwest Evidence-Based Practice Center, Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland, OR, United States

[2]Centers for Comparative and Health System Effectiveness, University of Washington, Seattle, WA, United States

[3]Veteran's Affairs Maine Healthcare System, Augusta, ME, United States

Corresponding Author:
William R Hersh, MD
Pacific Northwest Evidence-Based Practice Center
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
3181 SW Sam Jackson Park Rd.
BICC
Portland, OR,
United States
Phone: 1 503 494 4563
Fax: 1 503 494 4551
Email: hersh@ohsu.edu

Abstract

Background: Health information exchange (HIE), the electronic sharing of clinical information across the boundaries of health care organizations, has been promoted to improve the efficiency, cost-effectiveness, quality, and safety of health care delivery.

Objective: To systematically review the available research on HIE outcomes and analyze future research needs.

Methods: Data sources included citations from selected databases from January 1990 to February 2015. We included English-language studies of HIE in clinical or public health settings in any country. Data were extracted using dual review with adjudication of disagreements.

Results: We identified 34 studies on outcomes of HIE. No studies reported on clinical outcomes (eg, mortality and morbidity) or identified harms. Low-quality evidence generally finds that HIE reduces duplicative laboratory and radiology testing, emergency department costs, hospital admissions (less so for readmissions), and improves public health reporting, ambulatory quality of care, and disability claims processing. Most clinicians attributed positive changes in care coordination, communication, and knowledge about patients to HIE.

Conclusions: Although the evidence supports benefits of HIE in reducing the use of specific resources and improving the quality of care, the full impact of HIE on clinical outcomes and potential harms are inadequately studied. Future studies must address comprehensive questions, use more rigorous designs, and employ a standard for describing types of HIE.

Trial Registration: PROSPERO Registry No CRD42014013285; http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014013285 (Archived by WebCite at http://www.webcitation.org/6dZhqDM8t).

(JMIR Med Inform 2015;3(4):e39) doi:10.2196/medinform.5215

KEYWORDS

diagnostic tests; health information exchange; outcome assessment (health care); patient readmission; routine; systematic review

Introduction

In recent years, there has been substantial growth in the adoption of the electronic health record (EHR) in ambulatory and hospital settings across the United States, fueled largely by incentive funding provided by the Health Information Technology for Economic and Clinical Health (HITECH) Act. Following HITECH, 94% of nonfederal hospitals [1], 78% of hospital-based physicians [2], 84% of emergency departments (EDs), and 73% of hospital outpatient departments in the United States have adopted EHRs [3]. The motivation to increase the adoption of EHRs is grounded in evidence that health information technology (HIT) can improve the quality, safety, efficiency, and satisfaction with care, as has been reported in a series of systematic reviews [4-7].

One key challenge to effective use of HIT, however, is that most patients in the United States, especially those with multiple conditions, receive care across a number of settings [8,9]. To enable data to follow patients wherever they receive care, attention has recently focused on health information exchange (HIE), defined as the reliable and interoperable electronic sharing of clinical information among physicians, nurses, pharmacists, other health care providers, and patients across the boundaries of health care institutions, health data repositories, laboratories, public health agencies, and other entities that are not within a single organization or among affiliated providers [10].

The Office of the National Coordinator for Health Information Technology (ONC) has defined the following forms of HIE [11]:

1. *Directed exchange:* Sending and receiving secure information electronically between care providers.
2. *Query-based exchange:* Provider-initiated requests for information on a patient from other providers.
3. *Consumer-mediated exchange:* Patients aggregating and controlling the use of their health information among care providers.

ONC also uses the words “push” to describe directed exchange and “pull” to describe query-based exchange [12]. ONC leadership has also advocated that HIE be thought of as a verb and not as a noun, with more focus on the action of exchange and what is achieved with the information than on the technological and organizational structures required [13]. This is not meant to imply that the structures are not necessary, rather it is designed to shift the focus when evaluating HIE from documenting what has been created to the impact HIE has on health and health care.

The HITECH Act recognized that EHR adoption alone was insufficient to realize the full promise of HIT, allocating US $563 million for states or state-designated entities to establish HIE capability among health care providers and hospitals [11]. As a result of HITECH funding, HIE adoption has grown in a parallel though somewhat smaller manner. By 2014, 76% of US hospitals had engaged in some form of HIE [14]. An annual survey of organizations engaged in HIE found 135 in the United States in 2014 [15].

Evaluating the effectiveness of HIE (and HIT generally) has been challenging [16]. HIE systems are intermediate to improving care delivery, allowing clinicians and others improved access to patient data to inform decisions, and facilitate appropriate use of testing and treatment. HIE is not specific to any health issue or diagnosis. HIE implementations have often been supported by one-time start-up funding, without long-term support to sustain the programs long enough for evaluation.

There are 3 previously published systematic reviews that focus exclusively on HIE [17-19]. One of these reviews was conducted a half-decade ago [17], another focused only on US-based and clinical-only (ie, not public health) activities [18], and a third assessed mainly the associations between study characteristics and the frequency of positive outcomes [19]. We expanded upon these reviews to not only perform a systematic review of HIE but also determine needs for future research that reflect our assessment of the benefits and limitations of HIE.

Methods

Key questions guiding this review were developed by the review team with input from a group of stakeholders and the Agency for Healthcare Research and Quality (AHRQ). A standard protocol was developed using input from key informants and a technical expert panel, registered in PROSPERO [20], and posted on an AHRQ public website. A technical report further describes the methods and includes search strategies and additional information [21]. A research librarian conducted electronic database searches identifying relevant articles published between January 1990 and February 2015 in MEDLINE (Ovid), PsycINFO, CINAHL, and the Cochrane Library databases. Searches were peer reviewed by another librarian and supplemented by references identified from additional sources, including reference lists, table of contents of journals not indexed in databases searched, gray literature sources, and experts. English-language studies of HIE that reported on clinical, economic, population, and intermediate (eg, patient or provider perceptions, availability or accuracy of data, or time saved) outcomes were included. We included comparative studies of effectiveness, and other designs for more qualitative outcomes. We excluded studies that investigated benefits of HIE other than in clinical or public health settings (eg, to enhance clinical research). Two investigators independently evaluated each study to determine inclusion eligibility. Disagreement was resolved by consensus with a third investigator making the final decision as needed.

Details of included studies were extracted by one investigator and reviewed for accuracy and completeness by a second investigator. Two investigators independently assessed risk of bias for all effectiveness studies. Differences were resolved by discussion and consensus and reviewed by the team of investigators. Individual studies were rated as “low,” “moderate,” or “high” risk of bias. Investigators then assessed the strength of the body of evidence. Both the risk of bias and strength of evidence ratings were conducted using the criteria and procedures described in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews [22].

The strength of evidence consisted of the following 4 major categories: high, moderate, low, or insufficient, based on the methodological limitations of studies; consistency across studies; precision of estimates; and directness of effect. Ratings were reviewed by a second investigator, and disagreements were resolved by consensus or involvement of a third investigator if necessary. Data could not be combined in a quantitative meta-analysis because of heterogeneity in the interventions, the outcomes measured, and the way data were reported. Therefore, we combined studies qualitatively based on the similarity of the type of HIE, the implementation of the HIE, outcomes measured, and results reported. Where studies were not similar in these areas, we provided the results of the individual studies without grouping them.

Results

Of the 5211 potentially relevant citations identified in our literature searches, 849 articles were selected for full-text review and 34 studies were ultimately deemed to address outcomes. Study characteristics, results, and risk of bias assessments are presented in Multimedia Appendices 1 and 2. Of the studies included in this report, 2 were randomized controlled trials (RCTs) described in 3 papers and 32 were observational and survey studies. Most were conducted in the United States, although 8 were from Europe, Canada, Israel, and South Korea. These studies reported clinical or public health process, economic, or population outcomes; however, none of the studies explicitly stated that they assessed for harms of HIE or reported any negative unintended consequences. The majority were assessed to be of low risk of bias (ie, good internal validity) but also contained mostly retrospective observational evidence.

Of 34 studies, 26 reported clinical, economic, or population outcomes (see Multimedia Appendix 1), whereas the other 8 were found to report on perceptions of outcomes (see Multimedia Appendix 2). None of the studies evaluated primary clinical outcomes from HIE (eg, mortality and morbidity) nor explicitly measured or reported harms. We list the study designs and geographic locations in Table 1.

The most common study design for assessing outcomes was retrospective cohort, typically with HIE use associated with a specific outcome (Table 1). The next most common design was survey, which was usually focused on perception of effectiveness and perceived outcomes: 2 studies were RCTs—1 RCT assessed a particular directed information exchange (2 published papers, 1 on clinical outcomes, and 1 on perceptions) and the other evaluated a clinical decision support intervention using data from an HIE implementation. Two studies used cross-sectional analyses of large databases to compare health care organizations having access to HIE with those without access. Two other studies used a case series methodology, one of which involved asking clinicians if HIE access avoided undesirable resource use, and then calculating the costs saved and the other that retrospectively analyzed data to determine duplicative testing averted.

The identified studies were performed mostly in the United States, but we identified 8 studies from 5 other countries. Of the 26 studies in the United States, 2 assessed multiple HIE implementations across the entire United States, 1 assessed multiple HIE implementations in 2 states (California and Florida), and the remaining 23 studies were conducted in 13 states. Most studies used retrospective designs, usually with an approach examining the association of HIE use with 1 or more clinical variables. All of these studies focused on the direct effect of HIE, usually reporting reduction in resource use or costs, without determining its larger impact (eg, overall total or proportion of spending in an ED vs the total dollars that HIE appeared to save). None of the studies analyzed individual episodes of care to determine clinical appropriateness of possible changes brought about by HIE use.

The prospective studies also had limitations. The 2 RCTs (reported in 3 papers) were focused on highly specific uses of HIE, namely, directed exchange of ED reports in one and pharmacotherapy clinical decision support in another. Of note, however, was that neither study showed benefit of HIE. The other prospective study was a case series that was limited by its methodology relying on physician self-reports of resources not utilized when HIE was used, with no follow-up or validation of their decisions, or analysis of more holistic views of clinical outcomes or costs.

Table 1. Study designs and locations.

	Study designs and locations	References
Designs (number)		
	Retrospective cohort (18)	[23-40]
	Survey (8)	[41-48]
	Randomized controlled trial (2 reported in 3 papers)	[49-51]
	Cross sectional (2)	[52,53]
	Case series (2)	[54,55]
Location (number)		
	Austria (1)	[47]
	Canada (2)	[49,51]
	Finland (2)	[23,46]
	Israel (2)	[29,56]
	South Korea (1)	[48]
	All of United States (2)	[41,53]
	California and Florida	[52]
	Colorado (1)	[24]
	Indiana (3)	[35,36,44]
	Louisiana (1)	[34]
	Massachusetts (1)	[45]
	Minnesota (1)	[55]
	North Carolina (1)	[50]
	New York (6)	[32,33,37,40,42,43]
	Oklahoma (1)	[38]
	South Carolina (1)	[54]
	Tennessee (3)	[25,27,28]
	Texas (1)	[31]
	Virginia (1)	[39]
	Wisconsin (2)	[26,30]

Most of these studies had reasonable but not strong internal validity. As the intervention (HIE) was only one of many potential influences on clinical outcome (ie, many more factors go into clinical outcomes than the decision to consult an HIE on a patient), there was possible confounding. Because no confounders were explicitly identified and incorporated into the analyses, most studies with appropriate retrospective methods were rated as having low or moderate risk of bias.

Because of the type of study designs used, reporting limitations, and the lack of ability to combine results, the strength of this body of evidence was rated as low, meaning that future studies have the potential to alter these findings in magnitude or direction. In addition, the number of studies and their locations in the United States represent a small fraction of functioning HIE systems. A larger number are reported to be operational, sustainable, or innovating according to the eHealth Initiative Annual Data Exchange Survey, which reported a total of 84 such HIE implementations in 2013 [57] and 106 in 2014 [15]. In other words, while a substantial number of HIE implementations exist in the United States, only a small number have been subject to evaluation. This low number of studies relative to HIE efforts also makes it difficult to generalize about what aspects of HIE, such as location, type, and setting, are associated with the results reported in research.

Improving Resource Use

Most of the studies of HIE effectiveness focused on resource use. We categorized these as follows (Table 2): laboratory testing, radiology testing, hospital admissions, hospital readmissions, referrals and consultations, ED costs, public heath reporting, quality of care, and other aspects of HIE. Although the risk of bias in most studies was low to moderate, the resulting evidence from them was mostly of low strength due to retrospective designs. This low-strength evidence mostly favored the value of HIE in reducing resource use and costs, especially in the ED, but used a very narrow cost perspective

and did not account for how HIE was used and its impact on the overall care of the patient beyond the immediate setting where it was used.

Table 2. Study results by categories.

Category (number)	Results
Laboratory testing (6)	A total of 6 studies showed benefit for health information exchange (HIE) in reducing overall testing, although estimates of impact on cost were mixed [23-26,54,55]: 4 studies took place in the emergency department (ED) setting, all showing some amount of reduced testing and cost savings [25,26,54,55], whereas 2 studies were conducted in ambulatory settings, with one showing an increase [23] and the other showing a reduction in the increased overall rate of testing [24].
Radiology testing (9)	A total of 7 studies carried out in the ED setting showing reduced testing [25-28,52,54,55]; 2 studies were conducted in ambulatory settings, with one showing a decrease [23] and the other showing no change in the rate of testing [24].
Hospital admissions (8)	A total of 2 studies found a reduction in hospital admissions and lower costs [25,54]; 3 other studies also measured some benefit for HIE use in reducing hospital admissions [29,32,56], although 3 additional studies found no such reduction [30,31,49].
Hospital readmissions (2)	Whereas 1 study showed benefit for HIE in reducing hospital readmissions [33], the other did not [53].
Referrals and consultations (2)	A total of 2 studies assessed HIE for reducing referrals and/or consultations, with conflicting results [23,54].
ED costs (2)	A total of 2 studies found reduced overall ED costs per patient when HIE was available [25,26]. Neither study reported overall ED expenditures, making it unknown what proportion of overall ED spending was impacted by HIE.
Public heath reporting (3)	A total of 3 studies assessed HIE in public health settings, all of which were conducted in the United States and reported improved automated laboratory reporting [36], improved completeness of reporting for notifiable diseases [35], and improved identification of HIV patients for follow-up care [34].
Quality of care in ambulatory settings (3)	A total of 2 retrospective studies found HIE associated with improved quality of care [37,38], whereas a randomized controlled trial focused on medication reconciliation found increased ability to detect medication adherence problems, the results did not show improvement in adherence after it was identified and addressed by providers [50].
Other aspects of HIE (3)	A total of 3 studies assessed other aspects of HIE, including reduction in time for processing of Social Security Disability claims [39], increased ability to identify frequent ED users [40], and associated HIE implementation with improved patient satisfaction scores in hospitals [41].

Perceptions

A number of studies evaluated clinician or patient perceptions of outcomes of HIE (see Multimedia Appendix 2), with all reporting perceptions that HIE leads to some benefit including improved outcomes. Clinician perceptions of the value of HIE, where studied, were generally positive. However, how such perceptions translate into improved care is unknown. This body of evidence was considered low strength.

Factors Associated With Outcomes

To determine whether effectiveness of HIE varied by study type, health care setting, location, or HIE type, we examined whether HIE was found to have some beneficial effect or not across characteristics. As presented in Table 1, the preponderance of studies reported that HIE use for different functions, in various settings, and of varying types produced mostly positive outcomes. Although the number of positive versus negative studies was not an indicator of the overall direction of the evidence, we did note that for each “negative” study, there was at least one “positive” one. For type of HIE, there was no clear pattern of findings to suggest that one type was clearly better than another, even indirectly. The 2 RCTs reported no benefit for their selected outcomes from HIE intervention [49,50], although a perceptions study from one of them reported impressions of improved patient outcomes and management [51]. These were in contrast to the observational study designs where almost all found beneficial effects of HIE. For the HIE setting, only ambulatory and ED had enough studies to evaluate patterns, with outpatient settings less likely to find beneficial results compared with studies in ED settings. The sparseness of studies across geographic settings did not allow for identification of patterns, although across most studies in the United States, the findings were positive.

Discussion

A collection of low-quality evidence supports the value of HIE for reducing duplicative laboratory and radiology test ordering, lowering ED costs, reducing hospital admissions (less so for readmissions), improving public health reporting, increasing ambulatory quality of care, and improving disability claims processing. The evidence is low quality because of the retrospective nature of the studies and the limited questions that they ask. It is unlikely that additional studies of the kind included in this review will advance the field and strengthen our understanding when HIE can reduce laboratory and imaging tests associated with episodes of care without broadening their scope and using more rigorous designs. Although the preponderance of evidence reports positive effects of HIE in reducing resource use and improving quality of care, it is entirely possible that focused studies with stronger study designs and more comprehensive assessment of utilization or clinical outcomes might reach a different conclusion.

We found no studies explicitly addressing patient-specific clinical outcomes such as morbidity, mortality, or functional

status, and therefore the body of evidence is insufficient to determine whether HIE has an impact on patient outcomes. We also did not identify any studies that used systematic and comprehensive economic analysis. Although some of the studies we included projected or estimated cost savings based on measured changes in utilization or perceptions of clinicians, there were no studies that explicitly measured costs and assessed economic impact in a comprehensive fashion. It is fair to say, then, that there was insufficient evidence to reach conclusions on the economic impact of HIE.

Applicability

How likely are the effects reported in this review to be observed when applied under diverse conditions in health systems, hospitals, and clinics in the United States? The greatest confidence in the applicability of these findings comes from the breadth of settings—geographic, organizational, and technical—from which they are derived. By contrast, there are limitations to the applicability of the findings (beyond limitations to the internal validity already mentioned) having to do with these main concerns: (1) concentration of evidence from a relatively small number of HIE systems; (2) use of internally developed and refined health IT systems compared with local instances of commercial systems; and (3) the exceptionally broad variety of systems, contexts, and purposes of HIE reported in the studies included in this review.

First, the concern that the bulk of the evidence about health IT impact arises out of a relatively small number of centers has been raised before [4]. These centers have been referred to as "health IT leaders," which are typically large academic medical centers with internally developed health IT systems, implemented incrementally, and refined over a long period. The nature of the health IT systems is in each case unique (being locally developed), and more importantly it is difficult to separate the effects of the health IT from the confounding influences of the health system itself. However, whether findings from these systems can be generalized to the very different context of health system and hospital implementations of commercially developed systems over shorter periods with less internal development and implementation infrastructure has been called into question [4]. This "health IT leader" effect appears to be reduced in more recent updates to the 2006 systematic review by Chaudhry et al [4] but the issue remains important [5,7]. In this review of HIE, the concentration of evidence phenomenon is also present, with large numbers of published studies emanating from relatively few areas, this time regional implementation programs rather than academic health centers, such as Texas, New York, and the MidSouth e-Health Alliance.

Second, separate from the "health IT leader" concern, which has to do with the organizational capacity, resources, and mission of these centers, is the issue of internally developed systems compared with commercially developed systems. Although few of the studies we included described whether their software used was commercial or locally developed, the overall model of health IT purchase and installation of nonhealth IT leaders are usually quite different from that of the incremental internal development, implementation, and refinement that are seen in systems such as the Department of Veterans Affairs or the aforementioned "health IT leader" systems. Related to this concern is a finding from other aspects of health IT [58], namely, clinical decision support, where systems evaluated by their developers tend to achieve more positive outcomes from their evaluation than external evaluators. This phenomenon must be assessed with HIE as well.

Third, and most important, in terms of limiting the applicability of these findings about HIE to real-world use is the exceptionally wide variety of systems, purposes, and contexts of use. To predict whether specific implementations of HIE in specific health care contexts will have favorable impacts on specific desired outcomes is not possible from this review and in most cases would not be possible from comparison with individual studies because (1) it is unlikely that studies with low risk of bias have been published for most such specific questions, and (2) in almost all cases these are complex interventions that are incompletely specified, with insufficient detail to draw strong meaningful inferences [59].

Limitations of the Evidence Base

The significant limitations of the evidence base, that is, the individual studies included in this review, have been raised in previous systematic reviews of health IT [4,5,7] and of HIE [18]. There are four primary concerns about the limitations of the available evidence on the impact of HIE (and health IT in general): (1) suitability of study design; (2) execution of the studies; (3) complexity of the interventions with implications for interpretation and for generalizability; and (4) changes in the technology or policy governing its use.

First, the evidence in this area addresses a wide variety of questions covering diverse domains beyond medical science from computer science, human factors, sociology, organization and management, and other disciplines. This broad array of questions calls for an equally diverse range of study designs. Studies of usability and use require usability engineering methods, studies of individual behavior call for methods from anthropology and behavioral sciences, studies of organizational change warrant methods drawn from management and systems science, whereas studies of population effects call for the methods of epidemiologists. A significant limitation of this literature, with its breadth of research questions, is the limited toolbox often drawn upon to answer them.

The second limitation is in execution of the studies. Even when strong study designs are used, their execution may be lacking, whether in sampling strategies, measurement methods, or analytic approaches. The unit of analysis problem is but one example. Interventions carried out at the level of the health system, hospital, or clinic may be analyzed at the level of the patient or episode, without controlling for variation at these multiple levels. Incomplete measurement is another: for example, where ED test ordering is measured in isolation, ignoring the possibility that the same test might later be ordered in another setting such as urgent care, primary care, or in hospital.

The third limitation has to do with the complexity of interventions, where the HIE or other health IT system itself is

necessarily only part of a more complex intervention. The complexity of interventions to change the behavior of clinicians or others in the health systems studied requires more thorough specification, to both adjust for confounders and make sense out of how to apply interventions elsewhere. Others have documented the inadequacy of specification of the details of complex interventions and called for a more systematic and thorough reporting [59,60].

Finally, the literature does not comprehensively cover changes in technology or policies governing its use. For example, whereas most studies come from the locally developed systems of HIE leaders as noted earlier, there has been a more recent growth in the commercial marketplace for HIE. In addition, the widespread adoption of EHRs under the HITECH Act in the US means that a more diverse array of health care organizations will be participating in HIE implementations. As an example of policy changes governing HIE development, as noted in Table 1, most studies have been of query-based systems whereas more recent meaningful use criteria for incentive funding call for implementation of directed exchange.

Future Research Needs

Given the limited conclusions that can be reached after review of the large volume of published literature on HIE, what are the implications for future research? Recognizing that HIE, like health IT in general, will almost certainly undergo increasingly widespread implementation in the future, the first aim of researchers should be to shift the emphasis from *whether* HIE systems should be implemented to specifically *how* they should be implemented. The question to be answered is not "Does HIE have positive effects?" but rather "How can HIE be implemented in order to result in the greatest benefit for patients, clinicians, and health systems with the least cost and harm?"

A second aim of research on HIE should be to develop greater focus and clarity about the level at which interventions are operating and the types and levels at which outcomes are measured. The outcomes of interest and the factors influencing them may be quite different at different levels of analysis, from specific systems or functionalities of HIE to individual patients, providers, or episodes of care; to health care units such as the ED, primary care practice, or hospital ward; to institutions such as hospitals; to aggregates such as health systems; or to broader regional multiorganization entities or regions. Combining or confusing these levels of intervention and levels of analysis only increases the challenges for those who conduct the research and for those who wish to interpret and apply it.

To help achieve an improved focus and clarity, a more formal analytic framework and a more descriptive taxonomy are needed. An example of such a framework that could be usefully applied in this area is Rasmussen's sociotechnical hierarchy, which specifies the multiple levels of a complex sociotechnical system that must be considered together to understand system behavior change [61]. Examples of its application include Vicente's analysis of the forces acting at multiple levels to reduce hazards arising from patient-controlled analgesia devices [62] and Leveson's Systems—Theoretic Accident Modeling and Processes model for understanding system performance and safety [63].

Similarly, a formal taxonomy for implementation of complex interventions has been proposed that would enable more complete and useful specification of interventions to allow better analysis, interpretation, and application [59,64]. This taxonomy should be extended specific to HIE to include clinical, technical, and organizational details of the HIE implementation as outlined by Vest [65]. The clinical taxonomy should focus not only on patient outcomes, but also on issues such as health disparities related to HIE and health system issues that may improve or undermine use of HIE. The technical taxonomy should include aspects of system architecture, messaging and terminology standards, and other details. It should also address the financial aspects of implementations, such as whether locally developed or commercial software is used and whether the HIE organization is public or private. The HIE research community should consider a standardized reporting instrument for HIE evaluation comparable to the Consolidated Standards of Reporting Trials statement for RCTs [66].

The third step researchers can take to improve the evidence base for implementation of HIE is to broaden the methodologic toolbox applied to these questions. As indicated earlier, the study approach and architecture must be suited to the question being asked, employing methods from usability engineering, behavioral sciences, systems engineering, and organizational sciences, depending on the question being addressed. These would include methods used in engineering and quality improvement, as well as in the study of complex adaptive systems.

What types of studies should be performed? RCTs are impractical for technologies with wide-ranging purposes like HIE. Yet, retrospective studies associating HIE versus nonuse for outcomes such as test ordering and hospital admissions are very limited in conclusions that can be drawn. Research is also challenging because many of the important clinical outcomes that could be positively affected by HIE have many other potential contributing and confounding factors relating to the patient, his or her clinicians, the quality of care delivered, the EHR, other health IT used, the nature of the health care delivery system, and the regulatory environment. Given the growing evidence based on robust evaluations in other areas of health IT, as noted in systematic reviews [7], methodological insights can be gleaned from other topic areas.

Future studies should be prospective, carried out in mature HIE settings, specify a priori what patients and/or use cases are likely to benefit from HIE, and compare appropriate outcomes for the use or nonuse of HIE. The prospective collection of data from diverse settings where HIE is used, classified by the taxonomy advocated earlier, could allow for prospective cohort studies that could identify aspects of HIE associated with beneficial outcomes. This will likely require an effort comparable in scope to national data collection efforts, such as the Patient-Centered Outcomes Research Institute Clinical Data Research Network initiative [67]. Ideally, such an undertaking could be synergistic with these other large-scale efforts.

Evaluation should be a requirement for all HIE implementations, certainly those funded by grants or other external funding. The challenge of evaluating health IT projects, especially in

community settings, is well-known [16], but all funders must demand this requirement to grow the evidence base. By the same token, funders must provide adequate resources for such evaluations. In addition, evaluations should be performed by researchers external to the project to reduce potential bias from system developers evaluating their own implementations [58].

Conclusions

The full impact of HIE on clinical outcomes and potential harms is insufficiently studied, although evidence provides some support for benefit in reducing use of some specific resources and achieving improvements in quality of care measures. To advance our understanding of HIE, future studies need to address comprehensive questions, use more rigorous designs, and be part of a coordinated, systematic approach to studying HIE. Going forward, HIE will become a more integrated part of health care delivery, and its evaluation needs to be focused on maximizing the improvements that HIE usage brings to overall clinical care.

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: Andrew Hamilton, MLS, MS, for conducting literature searches and Spencer Dandy, BS, for assistance with preparing this manuscript (both are located at the Oregon Health & Science University); and Jon White, MD, Task Order Officer at the Agency for Healthcare Research and Quality (AHRQ). The primary funding source was The Agency for Healthcare Research and Quality (Contract Number 290-2012-00014-I, Task Order 11), Rockville, Maryland. SZK was also supported by a grant from the National Library of Medicine Training (Grant No T15LM007088). The findings and conclusions in this document are those of the authors, who are responsible for its content, and do not necessarily represent the views of AHRQ or the National Library of Medicine. No statement in this report should be construed as an official position of AHRQ or the US Department of Health and Human Services.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Studies of health information exchange included for assessing outcomes.

[PDF File (Adobe PDF File), 115KB - medinform_v3i4e39_app1.pdf]

Multimedia Appendix 2

Patient and clinician survey perceptions of health information exchange.

[PDF File (Adobe PDF File), 19KB - medinform_v3i4e39_app2.pdf]

References

1. Dustin C, Meghan G, Furukawa M. ONC Data Brief: Number 16. 2014. Adoption of electronic health record systems among US non-federal acute care hospitals: 2008-2013 URL: https://www.healthit.gov/sites/default/files/oncdatabrief16.pdf [accessed 2015-10-07] [WebCite Cache ID 6c6lYtrUT]
2. Hsiao C, Hing E. NCHS Data Brief Number 143. 2014. Use and characteristics of electronic health record systems among office-based physician practices: United States, 2001-2013 URL: http://www.cdc.gov/nchs/data/databriefs/db143.htm [accessed 2015-10-07] [WebCite Cache ID 6c6lpDgb3]
3. Eric J, Esther H. NCHS Data Brief Number 187. 2015. Progress with electronic health record adoption among emergency and outpatient departments: United States, 2006-2011 URL: http://www.cdc.gov/nchs/data/databriefs/db187.htm [accessed 2015-10-07] [WebCite Cache ID 6c6lwBTBp]
4. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [Medline: 16702590]
5. Goldzweig CL, Towfigh A, Maglione M, Shekelle PG. Costs and benefits of health information technology: New trends from the literature. Health Aff (Millwood) 2009;28(2):w282-w293 [FREE Full text] [doi: 10.1377/hlthaff.28.2.w282] [Medline: 19174390]
6. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: A review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471 [FREE Full text] [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
7. Jones SS, Rudin RS, 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]

8. Bourgeois FC, Olson KL, Mandl KD. Patients treated at multiple acute health care facilities: Quantifying information fragmentation. Arch Intern Med 2010 Dec 13;170(22):1989-1995. [doi: 10.1001/archinternmed.2010.439] [Medline: 21149756]
9. Finnell JT, Overhage JM, Grannis S. All health care is not local: An evaluation of the distribution of emergency department care delivered in Indiana. AMIA Annu Symp Proc 2011;2011:409-416 [FREE Full text] [Medline: 22195094]
10. What is HIE (health information exchange)?. 2012. URL: https://www.healthit.gov/providers-professionals/health-information-exchange/what-hie [accessed 2015-10-07] [WebCite Cache ID 6c6m5qCuS]
11. Williams C, Mostashari F, Mertz K, Hogin E, Atwal P. From the Office of the National Coordinator: The strategy for advancing the exchange of health information. Health Aff (Millwood) 2012 Mar;31(3):527-536 [FREE Full text] [doi: 10.1377/hlthaff.2011.1314] [Medline: 22392663]
12. Dullabh P, Ubri P, Loganathan S, Latterner M. Evaluation of the State Health Information Exchange Cooperative Agreement Program: State Approaches to Enabling HIE: Typology Brief. Washington, DC: Office of the National Coordinator for Health Information Technology; 2014. URL: https://www.healthit.gov/sites/default/files/statehietypologybrief.pdf [accessed 2015-10-07] [WebCite Cache ID 6c6obU0FC]
13. Healthcare Information and Management Systems Society. Healthcare IT News. 2015. URL: http://www.healthcareitnews.com/news/hie-verb-onc-wants-move-quickly-data-exchange [accessed 2015-10-07] [WebCite Cache ID 6c6mPt5aH]
14. Report N. Health Information Exchange Among US Non-Federal Acute Care Hospitals: 2008-2014 (ONC Data Brief Number 24). Washington, DC: Office of the National Coordinator for Health Information Technology (ONC); 2015. URL: https://www.healthit.gov/sites/default/files/data-brief/ONC_DataBrief24_HIE_Final.pdf [accessed 2015-10-07] [WebCite Cache ID 6c6okXYub]
15. eHealth I. Post HITECH: The Landscape of Health Information Exchange. 2014 Sep 10. URL: http://visit.medicity.com/rs/aetnainc/images/2014%20eHI%20Data%20Exchange%20Survey%20Key%20Findings.pdf [accessed 2015-10-07] [WebCite Cache ID 6c6mVWlpt]
16. Kern LM, Ancker JS, Abramson E, Patel V, Dhopeshwarkar RV, Kaushal R. Evaluating health information technology in community-based settings: Lessons learned. J Am Med Inform Assoc 2011;18(6):749-753 [FREE Full text] [doi: 10.1136/amiajnl-2011-000249] [Medline: 21807649]
17. Hincapie A, Warholak T. The impact of health information exchange on health outcomes. Appl Clin Inform 2011;2(4):499-507 [FREE Full text] [doi: 10.4338/ACI-2011-05-R-0027] [Medline: 23616891]
18. Rudin RS, Motala A, Goldzweig CL, Shekelle PG. Usage and effect of health information exchange: A systematic review. Ann Intern Med 2014 Dec 2;161(11):803-811. [doi: 10.7326/M14-0877] [Medline: 25437408]
19. Rahurkar S, Vest JR, Menachemi N. Despite the spread of health information exchange, there is little evidence of its impact on cost, use, and quality of care. Health Aff (Millwood) 2015 Mar;34(3):477-483. [doi: 10.1377/hlthaff.2014.0729] [Medline: 25732499]
20. Hersh W, Totten A, Gorman P, Eden K, Devine B, Woods S, et al. Health Information Exchange PROSPERO 2014:Crd42014013285. 2014. URL: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014013285 [accessed 2015-10-07] [WebCite Cache ID 6c6meV9wz]
21. Hersh W, Totten A, Eden K, Devine B, Gorman P, Kassakian S, et al. Health Information Exchange. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
22. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews: AHRQ Publication Number 10(14)-EHC062-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014.
23. Mäenpää T, Asikainen P, Gissler M, Siponen K, Maass M, Saranto K, et al. Outcomes assessment of the regional health information exchange: A five-year follow-up study. Methods Inf Med 2011;50(4):308-318. [doi: 10.3414/ME10-01-0048] [Medline: 21336419]
24. Ross SE, Radcliff TA, Leblanc WG, Dickinson LM, Libby AM, Nease DE. Effects of health information exchange adoption on ambulatory testing rates. J Am Med Inform Assoc 2013;20(6):1137-1142 [FREE Full text] [doi: 10.1136/amiajnl-2012-001608] [Medline: 23698257]
25. Frisse ME, Johnson KB, Nian H, Davison CL, Gadd CS, Unertl KM, et al. The financial impact of health information exchange on emergency department care. J Am Med Inform Assoc 2012;19(3):328-333 [FREE Full text] [doi: 10.1136/amiajnl-2011-000394] [Medline: 22058169]
26. Tzeel A, Lawnicki V, Pemble KR. The business case for payer support of a community-based health information exchange: A humana pilot evaluating its effectiveness in cost control for plan members seeking emergency department care. Am Health Drug Benefits 2011 Jul;4(4):207-216 [FREE Full text] [Medline: 25126351]
27. Bailey JE, Wan JY, Mabry LM, Landy SH, Pope RA, Waters TM, et al. Does health information exchange reduce unnecessary neuroimaging and improve quality of headache care in the emergency department? J Gen Intern Med 2013 Feb;28(2):176-183 [FREE Full text] [doi: 10.1007/s11606-012-2092-7] [Medline: 22648609]
28. Bailey JE, Pope RA, Elliott EC, Wan JY, Waters TM, Frisse ME. Health information exchange reduces repeated diagnostic imaging for back pain. Ann Emerg Med 2013 Jul;62(1):16-24. [doi: 10.1016/j.annemergmed.2013.01.006] [Medline: 23465552]

29. Ben-Assuli O, Shabtai I, Leshno M. The impact of EHR and HIE on reducing avoidable admissions: Controlling main differential diagnoses. BMC Med Inform Decis Mak 2013;13:49 [FREE Full text] [doi: 10.1186/1472-6947-13-49] [Medline: 23594488]
30. Tzeel A, Lawnicki V, Pemble KR. "Hidden" value: How indirect benefits of health information exchange further promote sustainability. Am Health Drug Benefits 2012 Sep;5(6):333-341 [FREE Full text] [Medline: 24991331]
31. Vest JR. Health information exchange and healthcare utilization. J Med Syst 2009 Jun;33(3):223-231. [Medline: 19408456]
32. Vest JR, Kern LM, Campion TR, Silver MD, Kaushal R. Association between use of a health information exchange system and hospital admissions. Appl Clin Inform 2014;5(1):219-231 [FREE Full text] [doi: 10.4338/ACI-2013-10-RA-0083] [Medline: 24734135]
33. Vest JR, Kern LM, Silver MD, Kaushal R. The potential for community-based health information exchange systems to reduce hospital readmissions. J Am Med Inform Assoc 2015 Mar;22(2):435-442. [doi: 10.1136/amiajnl-2014-002760] [Medline: 25100447]
34. Magnus M, Herwehe J, Gruber D, Wilbright W, Shepard E, Abrams A, et al. Improved HIV-related outcomes associated with implementation of a novel public health information exchange. Int J Med Inform 2012 Oct;81(10):e30-e38. [doi: 10.1016/j.ijmedinf.2012.06.005] [Medline: 22883431]
35. Dixon BE, McGowan JJ, Grannis SJ. Electronic laboratory data quality and the value of a health information exchange to support public health reporting processes. AMIA Annu Symp Proc 2011;2011:322-330 [FREE Full text] [Medline: 22195084]
36. Overhage JM, Grannis S, McDonald CJ. A comparison of the completeness and timeliness of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. Am J Public Health 2008 Feb;98(2):344-350. [doi: 10.2105/AJPH.2006.092700] [Medline: 18172157]
37. Kern LM, Barrón Y, Dhopeshwarkar RV, Kaushal R. Health information exchange and ambulatory quality of care. Appl Clin Inform 2012;3(2):197-209 [FREE Full text] [doi: 10.4338/ACI-2012-02-RA-0005] [Medline: 23646072]
38. Nagykaldi ZJ, Yeaman B, Jones M, Mold JW, Scheid DC. HIE-i-health information exchange with intelligence. J Ambul Care Manage 2014;37(1):20-31. [doi: 10.1097/JAC.0000000000000002] [Medline: 24309392]
39. Feldman SS, Horan TA. Collaboration in electronic medical evidence development: A case study of the Social Security Administration's MEGAHIT System. Int J Med Inform 2011 Aug;80(8):e127-e140. [doi: 10.1016/j.ijmedinf.2011.01.012] [Medline: 21333588]
40. Shapiro JS, Johnson SA, Angiollilo J, Fleischman W, Onyile A, Kuperman G. Health information exchange improves identification of frequent emergency department users. Health Aff (Millwood) 2013 Dec;32(12):2193-2198. [doi: 10.1377/hlthaff.2013.0167] [Medline: 24301405]
41. Vest JR, Miller TR. The association between health information exchange and measures of patient satisfaction. Appl Clin Inform 2011;2(4):447-459 [FREE Full text] [doi: 10.4338/ACI-2011-06-RA-0040] [Medline: 23616887]
42. Altman R, Shapiro JS, Moore T, Kuperman GJ. Notifications of hospital events to outpatient clinicians using health information exchange: A post-implementation survey. Inform Prim Care 2012;20(4):249-255. [Medline: 23890336]
43. Campion TR, Ancker JS, Edwards AM, Patel VN, Kaushal R. Push and pull: Physician usage of and satisfaction with health information exchange. AMIA Annu Symp Proc 2012;2012:77-84 [FREE Full text] [Medline: 23304275]
44. Chang KC, Overhage JM, Hui SL, Were MC. Enhancing laboratory report contents to improve outpatient management of test results. J Am Med Inform Assoc 2010;17(1):99-103 [FREE Full text] [doi: 10.1197/jamia.M3391] [Medline: 20064809]
45. Kaushal R, Dhopeshwarkar R, Gottlieb L, Jordan H. User experiences with pharmacy benefit manager data at the point of care. J Eval Clin Pract 2010 Dec;16(6):1076-1080. [doi: 10.1111/j.1365-2753.2009.01255.x] [Medline: 20666888]
46. Maass MC, Asikainen P, Mäenpää T, Wanne O, Suominen T. Usefulness of a regional health care information system in primary care: A case study. Comput Methods Programs Biomed 2008 Aug;91(2):175-181. [doi: 10.1016/j.cmpb.2008.04.005] [Medline: 18514363]
47. Machan C, Ammenwerth E, Schabetsberger T. Evaluation of the electronic transmission of medical findings from hospitals to practitioners by triangulation. Methods Inf Med 2006;45(2):225-233. [doi: 10.1267/METH06020225] [Medline: 16538293]
48. Park H, Lee S, Kim Y, Heo E, Lee J, Park JH, et al. Patients' perceptions of a health information exchange: A pilot program in South Korea. Int J Med Inform 2013 Feb;82(2):98-107. [doi: 10.1016/j.ijmedinf.2012.05.001] [Medline: 22658777]
49. Lang E, Afilalo M, Vandal AC, Boivin J, Xue X, Colacone A, et al. Impact of an electronic link between the emergency department and family physicians: A randomized controlled trial. CMAJ 2006 Jan 31;174(3):313-318 [FREE Full text] [doi: 10.1503/cmaj.050698] [Medline: 16399880]
50. Willis JM, Edwards R, Anstrom KJ, Johnson FS, Del Fiol G, Kawamoto K, et al. Decision support for evidence-based pharmacotherapy detects adherence problems but does not impact medication use. Stud Health Technol Inform 2013;183:116-125. [Medline: 23388267]
51. Afilalo M, Lang E, Léger R, Xue X, Colacone A, Soucy N, et al. Impact of a standardized communication system on continuity of care between family physicians and the emergency department. CJEM 2007 Mar;9(2):79-86. [Medline: 17391577]
52. Lammers EJ, Adler-Milstein J, Kocher KE. Does health information exchange reduce redundant imaging? Evidence from emergency departments. Med Care 2014 Mar;52(3):227-234. [doi: 10.1097/MLR.0000000000000067] [Medline: 24374414]

53. Jones SS, Friedberg MW, Schneider EC. Health information exchange, health information technology use, and hospital readmission rates. AMIA Annu Symp Proc 2011;2011:644-653 [FREE Full text] [Medline: 22195120]
54. Carr CM, Gilman CS, Krywko DM, Moore HE, Walker BJ, Saef SH. Observational study and estimate of cost savings from use of a health information exchange in an academic emergency department. J Emerg Med 2014 Feb;46(2):250-256. [doi: 10.1016/j.jemermed.2013.05.068] [Medline: 24071033]
55. Winden TJ, Boland LL, Frey NG, Satterlee PA, Hokanson JS. Care everywhere, a point-to-point HIE tool: Utilization and impact on patient care in the ED. Appl Clin Inform 2014;5(2):388-401 [FREE Full text] [doi: 10.4338/ACI-2013-12-RA-0100] [Medline: 25024756]
56. Ben-Assuli O, Shabtai I, Leshno M. Using electronic health record systems to optimize admission decisions: The Creatinine case study. Health Informatics J 2015 Mar;21(1):73-88. [doi: 10.1177/1460458213503646] [Medline: 24692078]
57. eHealth Initiative. Result from Survey on Health Data Exchange 2013. The Challenge to Connect. 2013. URL: https://ehi-rails-app.s3.amazonaws.com/uploads/article/file/3/eHIResultsFromSurveyonHealthDataExchange2013.pdf [accessed 2015-06-17] [WebCite Cache ID 6c6n4pAts]
58. Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, Van Spall Harriette G C, et al. Features of effective computerised clinical decision support systems: Meta-regression of 162 randomised trials. BMJ 2013;346:f657 [FREE Full text] [Medline: 23412440]
59. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: The need for a scientific method. Implement Sci 2009;4:40 [FREE Full text] [doi: 10.1186/1748-5908-4-40] [Medline: 19607700]
60. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: The new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
61. Rasmussen J. Risk management in a dynamic society: A modelling problem. Saf Sci 1997 Nov;27(2-3):183-213. [doi: 10.1016/S0925-7535(97)00052-0]
62. Vicente KJ. What does it take? A case study of radical change toward patient safety. Jt Comm J Qual Saf 2003 Nov;29(11):598-609. [Medline: 14619352]
63. Leveson N. Engineering a safer world. Systems thinking applied to safety. Cambridge: MIT Press; 2012. URL: https://mitpress.mit.edu/sites/default/files/titles/free_download/9780262016629_Engineering_a_Safer_World.pdf [WebCite Cache ID 6c6ovJCqs]
64. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
65. Vest JR. More than just a question of technology: Factors related to hospitals' adoption and implementation of health information exchange. Int J Med Inform 2010 Dec;79(12):797-806. [doi: 10.1016/j.ijmedinf.2010.09.003] [Medline: 20889370]
66. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 1;152(11):726-732. [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
67. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc 2014;21(4):578-582 [FREE Full text] [doi: 10.1136/amiajnl-2014-002747] [Medline: 24821743]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
ED: emergency department
EHR: electronic health record
HIE: health information exchange
HIT: health information technology
HITECH: Health Information Technology for Economic and Clinical Health
IT: information technology
ONC: Office of the National Coordinator for Health Information Technology
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 08.10.15; peer-reviewed by R Elmore, M Solomon; comments to author 06.11.15; revised version received 10.11.15; accepted 11.11.15; published 15.12.15.

Please cite as:
Hersh WR, Totten AM, Eden KB, Devine B, Gorman P, Kassakian SZ, Woods SS, Daeges M, Pappas M, McDonagh MS
Outcomes From Health Information Exchange: Systematic Review and Future Research Needs
JMIR Med Inform 2015;3(4):e39
URL: http://medinform.jmir.org/2015/4/e39/
doi:10.2196/medinform.5215
PMID:26678413

Original Paper

Technology for Large-Scale Translation of Clinical Practice Guidelines: A Pilot Study of the Performance of a Hybrid Human and Computer-Assisted Approach

Stijn Van de Velde[1,2*], PhD; Lieve Macken[3*], PhD; Koen Vanneste[4*]; Martine Goossens[1,2,5*], Lic. Roman Philology; Jan Vanschoenbeek[1*], MD; Bert Aertgeerts[1,2,5*], MD PhD; Klaar Vanopstal[3*], PhD; Robert Vander Stichele[1,6*], MD PhD; Joost Buysschaert[3*], PhD

[1]EBMPracticeNet, Leuven, Belgium

[2]Belgian Center for Evidence-Based Medicine, Leuven, Belgium

[3]Department of Translation, Interpreting and Communication, Ghent University, Ghent, Belgium

[4]IVS Iscientia, Keerbergen, Belgium

[5]Academic Centre for General Practice, Katholieke Universiteit Leuven, Leuven, Belgium

[6]Heymans Institute of Pharmacology, Ghent University, Ghent, Belgium

[*]all authors contributed equally

Corresponding Author:
Stijn Van de Velde, PhD
EBMPracticeNet
Kapucijnenvoer 33 Blok J
Leuven, 3000
Belgium
Phone: 32 16332697
Fax: 32 16332697
Email: stijn.vandevelde@med.kuleuven.be

Abstract

Background: The construction of EBMPracticeNet, a national electronic point-of-care information platform in Belgium, began in 2011 to optimize quality of care by promoting evidence-based decision making. The project involved, among other tasks, the translation of 940 EBM Guidelines of Duodecim Medical Publications from English into Dutch and French. Considering the scale of the translation process, it was decided to make use of computer-aided translation performed by certificated translators with limited expertise in medical translation. Our consortium used a hybrid approach, involving a human translator supported by a translation memory (using SDL Trados Studio), terminology recognition (using SDL MultiTerm terminology databases) from medical terminology databases, and support from online machine translation. This resulted in a validated translation memory, which is now in use for the translation of new and updated guidelines.

Objective: The objective of this experiment was to evaluate the performance of the hybrid human and computer-assisted approach in comparison with translation unsupported by translation memory and terminology recognition. A comparison was also made with the translation efficiency of an expert medical translator.

Methods: We conducted a pilot study in which two sets of 30 new and 30 updated guidelines were randomized to one of three groups. Comparable guidelines were translated (1) by certificated junior translators without medical specialization using the hybrid method, (2) by an experienced medical translator without this support, and (3) by the same junior translators without the support of the validated translation memory. A medical proofreader who was blinded for the translation procedure, evaluated the translated guidelines for acceptability and adequacy. Translation speed was measured by recording translation and post-editing time. The human translation edit rate was calculated as a metric to evaluate the quality of the translation. A further evaluation was made of translation acceptability and adequacy.

Results: The average number of words per guideline was 1195 and the mean total translation time was 100.2 minutes/1000 words. No meaningful differences were found in the translation speed for new guidelines. The translation of updated guidelines was 59 minutes/1000 words faster (95% CI 2-115; *P*=.044) in the computer-aided group. Revisions due to terminology accounted for one third of the overall revisions by the medical proofreader.

Conclusions: Use of the hybrid human and computer-aided translation by a non-expert translator makes the translation of updates of clinical practice guidelines faster and cheaper because of the benefits of translation memory. For the translation of new guidelines, there was no apparent benefit in comparison with the efficiency of translation unsupported by translation memory (whether by an expert or non-expert translator).

(JMIR Med Inform 2015;3(4):e33) doi:10.2196/medinform.4450

KEYWORDS

practice guidelines as topic; translations; technology; education, medical, continuing; evidence-based practice

Introduction

The construction of EBMPracticeNet, a national electronic point-of-care information platform for the Belgian context, was initiated in 2011 to optimize quality of care by promoting evidence-based decision making [1]. The fundamental principle of evidence-based medicine (EBM) is that diagnostic and therapeutic actions must be based on the best available scientific knowledge about possible decisions, supplemented with the clinical expertise of the provider and taking into account the values and preferences of the patient [2]. Evidence-based practice guidelines have been developed to help clinicians keep up to date with current evidence and to support the use of evidence-based medicine in practice. The Institute of Medicine defines guidelines as "Statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options."

Belgian scientific associations of primary care physicians have produced about 50 Belgian clinical practice guidelines of good quality linked to electronic health records through a tool called the Evidence Linker [1]. To be able to provide answers to a broader array of health questions posed by physicians at the point of care, we supplemented the set of guidelines with an international collection of evidence-based point-of-care summaries. Such point-of-care summaries have been defined as Web-based medical compendia specifically designed to deliver predigested, rapidly accessible, comprehensive, periodically updated, and evidence-based information to clinicians [3]. Based on a broad evaluation, we eventually chose to subscribe to the Duodecim Evidence-Based Medicine (EBM) Guidelines database with the intention to adapt them to the Belgian context [4]. This database contains (to date) 940 EBM Guidelines of Duodecim Medical Publications available in English.

The implementation of evidence-based information in a specific context is influenced by the interaction of determinants that can be grouped into 7 domains: guideline factors, individual health professional factors, patient factors, professional interactions, incentives and resources, capacity for organizational change, and social, political, and legal factors [5]. The accessibility of guidelines written in the mother tongue is an example of a specific factor that might influence an implementation strategy. Assuring the availability of international guidelines in the local language increases the chances that they are consulted by non-native English medical professionals and that the recommendations contained in them are better retained. Especially when guidelines are offered for point-of-care use by busy non-native English speaking physicians, translation in their mother language is essential for acceptance, ease of use, and adoption [6]. This point of view is based on limited but consistent evidence on the importance of language as a barrier to the use of evidence-based medicine [7-11]. The importance of translating English guidelines into Dutch and French was also pinpointed by the interviewees in a Belgian study on facilitating factors for the dissemination and implementation of guidelines [12]. It was for these reasons that the Belgian health care authorities ordered the translation of these Duodecim Guidelines from English into Dutch and French.

Considering the scale and potential cost of the translation project, it was decided to make use of computer-aided translation performed by certificated translators with limited expertise in medical translation. The output of the translations was subsequently revised by medical proofreaders (general practitioners).

The first cycle of 940 revised translations resulted in the construction of a validated medical domain-specific translation memory, possibly helpful for the translation of future new guidelines or future updates of existing guidelines. New Duodecim Guidelines, as well as updated existing guidelines, will continue to be translated using the same hybrid approach, supported by the translation memory. As this method involves human translation/validation as well as computer support, we will describe it as a "hybrid human and computer-assisted" approach. If the hybrid method were to approach the quality and speed of expert medical translation, the method may serve as a model to other major medical translation projects.

Therefore, this pilot study aims to compare the speed and quality of 3 approaches of translation: (1) a certificated junior translator without medical specialization using the hybrid method, (2) an experienced medical translator with medical specialization but without this support, and (3) a certificated junior translator without medical specialization without this support.

Methods

Construction of the Translation Memory

To construct a translation memory on the basis of the first 940 Duodecim Guidelines, junior translators (trained at the master's level in general English-French and English-Dutch, but without special training in medical translations) used a hybrid approach, involving human translation supported by translation memory and by terminology recognition from terminology databases (termbases), as well as support from online machine translation. The software used was SDL Trados Studio.

The output of the translations was subsequently revised by medical proofreaders (general practitioners). The resulting corrected versions of the first 940 translations have now been converted to a validated medical domain-specific translation memory.

The principle of translation memory is that it stores source and target segments during human translation and offers translation suggestions on the basis of earlier translations when an identical or similar segment is submitted for translation. Approximately 75,000 translated segments were generated during the preparatory period. In addition, terminology recognition is used. Any term in a submitted segment that is also present in a termbase attached to the translation project is immediately marked and its equivalent term in the target language is displayed. The memory is stored in the sdltm format of SDL Trados Studio. It is compatible with the sdltb format of the SDL MultiTerm termbases, used earlier in the project (approximately 5000 terms and their translations). SDL Trados Studio allows immediate segment-specific access to Google Translate, which uses statistical methods to suggest translations based on large bilingual corpora.

Further details on the development of the translation memory and on the termbases are described elsewhere [1].

Hybrid Method of Translation in the Experiment

In the experiment, junior translators used the same hybrid human and computer-aided approach, now relying on the full, validated translation memory as described earlier. The translators also made use of the termbases and Google Translate. The output was again validated by a medical proofreader.

The quality and speed of this method were compared with the quality and speed of 2 other translation methods based on human translation without the validated translation memory and without the termbases but allowing help from Google Translate.

Research Question

The research question was "What is the efficiency, measured in terms of the quality of the translation output and the speed of the translation procedures, of translation by (1) a certificated junior translator without medical specialization using the hybrid method (arm A), (2) an experienced medical translator with medical specialization but without the support of the translation memory and the termbases (arm B), and (3) a certificated junior translator without medical specialization without the support of the translation memory and the termbases (arm C)?"

Study Design

We conducted a three-armed study in which comparable guidelines were translated by the 3 described methods. Another approach could have been to translate the same guidelines by both A and B, but we decided not to do this because of costs.

We used stratified randomization according to the number of words and the Flesch Reading Ease formula to ensure that the 3 arms were as similar as possible [13]. The randomization was performed by a third person, who was not involved in this study. We used random number lists obtained from a randomization website [14]. Ethical committee approval was not required for this study.

Source Guidelines

We drew our sample from the Duodecim Evidence-Based Medicine Guidelines collection. These guidelines are targeted at primary care, ambulatory care, and community hospitals. The Duodecim guidelines are available in English and present recommendations in a concise way to increase the usability at the point of care. The collection is updated on a continuous basis, and every 3 years the entire collection is revised. New guidelines are published regularly. From this collection, we included all 30 new guidelines published between August 2011 and January 2014. In addition, we selected 30 guidelines with major updates, correcting or amplifying earlier information.

Hypothesis

The hypothesis was that the hybrid translation method, which allows reuse of earlier translations through its translation memory component, could substantially contribute to translation efficiency in the case of updated guidelines.

Translation Procedures

The procedures for each arm in the experiment are described in the following section and represented schematically in Figure 1.

In arm A, a certificated junior translator without medical specialization was asked to translate the English guidelines into Dutch using the hybrid approach explained earlier, that is, making use of the SDL Trados Studio translation software that provided input based on the translation memory (approximately 75,000 translated segments), the SDL MultiTerm termbases used earlier in the project (approximately 5000 terms and their translations), and support from Google Translate. The translator was also allowed to use other online and paper sources.

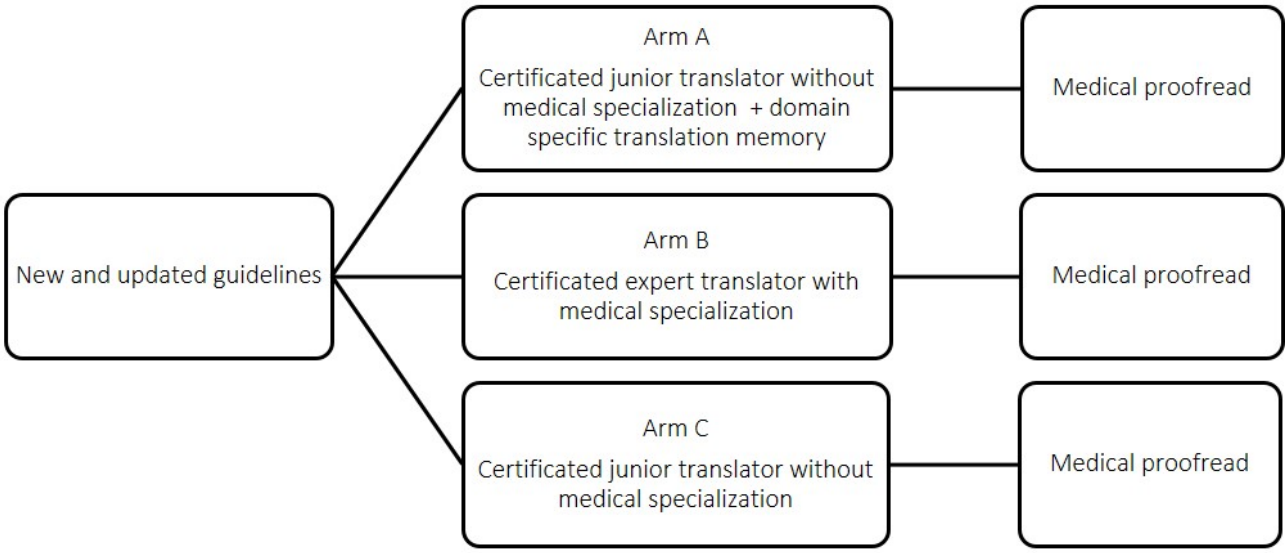
In arm B, an experienced medical translator (professional translator and editor at EBMPracticeNet) was allowed to use her own translation resources but did not have access to the validated translation memory or to the MultiTerm termbases. She was allowed to use input from Google Translate.

In arm C, the same certificated junior translator as in arm A translated guidelines to Dutch without access to the validated translation memory and to the MultiTerm termbases. The translator was allowed to use other sources, including Google Translate, that were deemed relevant but none were recommended to the translator by the team.

After translation, a medical proofreader, who was blinded to the translation procedure status, revised every translated guideline. Instructions for the medical proofreader included to repair nonsensical phrases; fix interpretation errors; rectify mistranslations, nontranslation, or inconsistent translation of terminology; and ensure that the text is understandable and stylistically acceptable to a Dutch native speaker who needs to understand the contents of the document [15]. The medical proofreader was also advised not to change text that is accurate and acceptable just for the sake of improving its style [16].

These procedures were applied once for the 30 new guidelines and once for 30 updated guidelines, but the junior translator was a different person in the second part of the experiment. By "certificated translator," we mean a translator with a university or college degree in translation.

Figure 1. Translation procedures.



Outcomes

Baseline Data

To evaluate the baseline comparability of the guidelines, for each guideline we counted the number of words and calculated the Flesch Reading Ease score [13]. The Flesch scale goes from 0 to 120. A lower score indicates more difficult text and a score below 30 is recommended only for a reader at the university graduate level. To calculate the score, we used Hendi, a tool developed to assess the readability of texts [17].

Translation Speed

Translation speed was measured by recording the translation time and the postediting time needed by the medical proofreader. For this purpose, we used Time Stamp [18]. To evaluate the speed of translation, we added up translation and revision time for each guideline.

Translation Quality

The output of the 3 types of translation was compared using the Human Translation Edit Rate (HTER). This is an automated metric based on edit distance that is usually used to calculate the minimum number of changes required for highly trained human editors to correct machine translation output so that it accurately reflects the meaning of the reference translation [19]. A higher HTER score indicates a higher number of changes. To calculate the HTER, we compared the output of every guideline after translation with the output after medical proofreading as the reference translation. The HTER was compared at the global text level.

To gain insights into the type of changes made by the medical proofreader, a further manual evaluation was made of translation acceptability and adequacy for 3 new and 3 updated guidelines per arm. Adequacy relates to the correspondence in meaning between source text and target text, whereas acceptability is the linguistic felicity of the target, that is, the use of suitable wording to express what was intended [20]. To evaluate adequacy and acceptability, we classified all revisions of the medical proofreader into 5 subcategories for adequacy and 5 subcategories for acceptability. See Figures 2 and 3 for an overview of the subcategories [21].

Figure 2. Number of revisions by the medical proofreader with respect to acceptability. Terminology refers to the use of other terms than those in the predefined list of preferred terms. Lexicon refers to bad word choice or use of wrong prepositions.

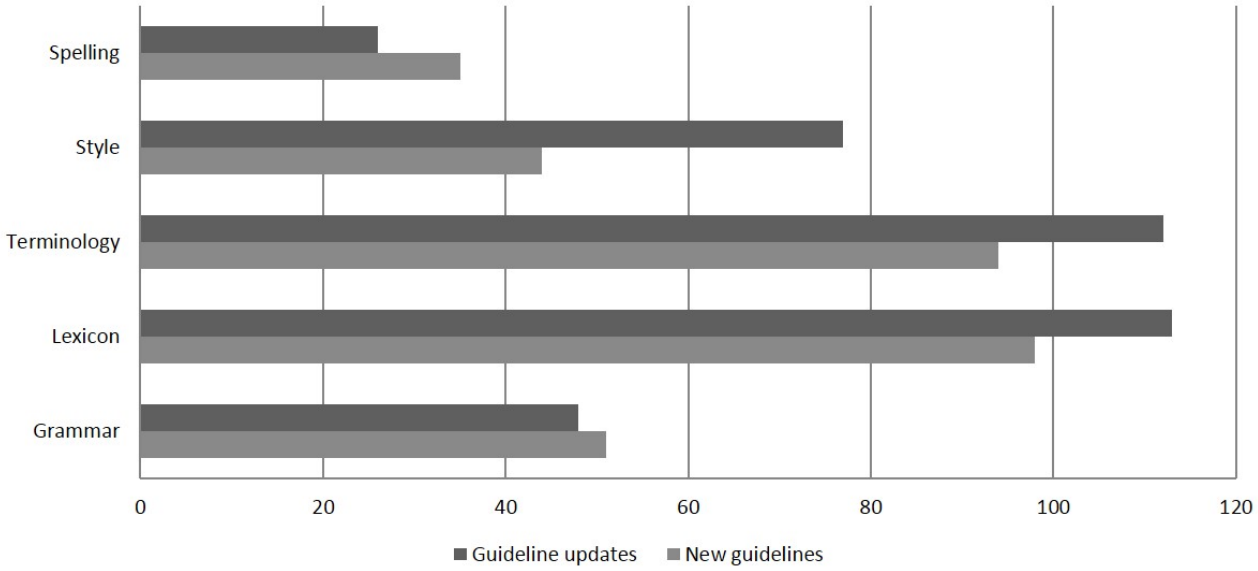
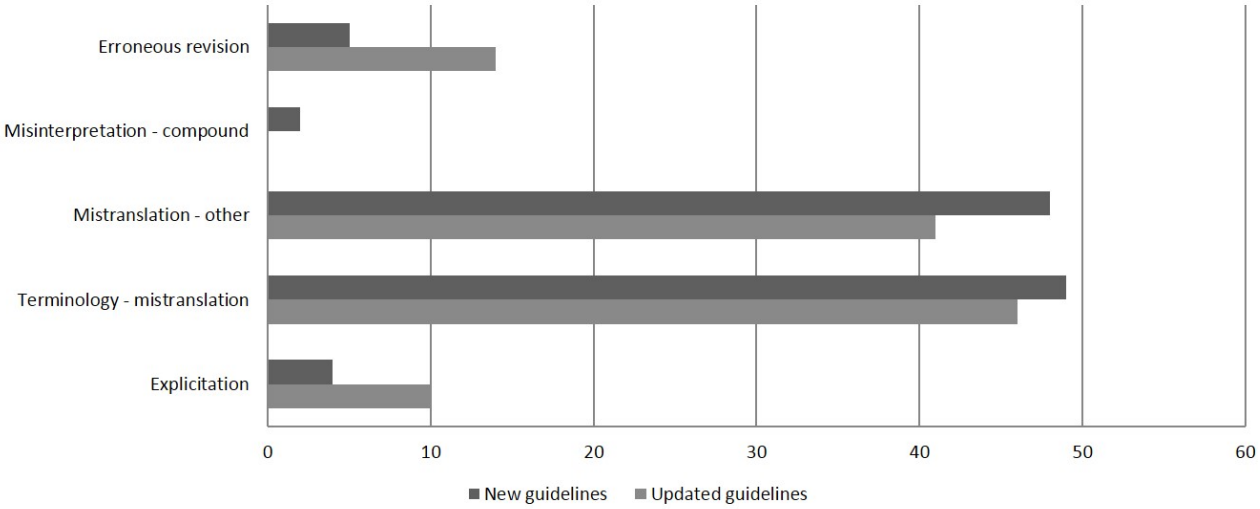


Figure 3. Number of revisions by the medical proofreader in relation to the adequacy domain. Explicitation means that the reviser amplified the translation to make its meaning more explicit. Terminology refers to the incorrect translation of terms, while Mistranslation refers to incorrect translation of other words. Misinterpretation indicates that a compound was misinterpreted.



Sample Size

Given that this was a pilot study with a fixed sample, we did not perform a priori sample size calculations.

Statistical Methods

We used descriptive statistics to compare baseline characteristics (number of words and Flesch score) of the 3 groups. The time needed per 1000 words was calculated, and these ratios were compared between groups using a one-way analysis of variance (ANOVA) followed by Tukey honest significant difference (HSD) tests. A *P* value less than .05 was considered significant. Given that it is difficult to evaluate the normality assumption in small datasets, nonparametric tests were conducted as well to verify the robustness of the drawn conclusions (sensitivity analysis). More specifically, a Kruskal-Wallis test followed by pairwise Mann-Whitney U tests were performed.

Results

Baseline Data

The average number of words per guideline was 1195 (new guidelines=1172; updated guidelines=1218). This illustrates the concise character of the point-of-care guidelines. Because we selected updates with major revisions, the number of changed or new words is very high. The Flesch reading score ranged from 31.4 to 36.6 corresponding to readability at a university graduate level. Table 1 shows the baseline data per study arm for the 30 new guidelines. Data per study arm for the updated guidelines are presented in Table 2. The number of words per guideline varied from 210 to 4695, which explains the large standard deviations reported in the tables.

Table 1. Baseline data for the new guidelines per comparison group.

Arm	Baseline data, mean (SD)				
	Words	FRE score	Translator (Time/1000w)	Proofreader (Time/1000w)	Total (Time/1000w)
A[a](n=10)	1252.4 (1334.8)	35.8 (10.3)	18.7 (4.1)	25.3 (9.8)	44.0 (12.2)
B[b](n=10)	1320.2 (991.5)	32.3 (10.4)	51.7 (13.5)	14.8 (3.4)	66.5 (15.6)
C[c](n=10)	943.8 (314.0)	36.1 (7.4)	22.6 (5.6)	23.4 (7.5)	45.9 (10.0)

[a]Certificated junior translator without medical specialization with domain-specific translation memory.

[b]Experienced medical translator without translation memory.

[c]Certificated junior translator without medical specialization without translation memory.

Table 2. Baseline data for the updated guidelines per comparison group.

Arm	Baseline data, mean (SD)					
	Words	Changed/new words	FRE score	Translator (Time/1000w)	Proofreader (Time/1000w)	Total (Time/1000w)
Arm A[a](n=10)	1376.0 (1211.1)	945.9 (501.0)	31.4 (5.9)	130.0 (50.5)	22.9 (14.2)	151.7 (59.8)
Arm B[b](n=10)	1100.0 (589.6)	1012.8 (642.0)	36.5 (5.5)	66.5 (20.6)	17.1 (8.3)	83.7 (27.1)
Arm C[c](n=10)	1178.3 (637.4)	1070.7 (681.9)	36.6 (7.2)	192.8 (53.2)	17.6 (7.9)	210.4 (57.5)

[a]Certificated junior translator without medical specialization with domain specific translation memory.

[b]Experienced medical translator without translation memory.

[c]Certificated junior translator without medical specialization without translation memory.

Translation Speed

Overall, the mean total translation time was 100.2 (70.7) minutes/1000 words. Translation by the experienced medical translator and medical proofreader took 66.5 min/1000 words for the new guidelines and 83.7 minutes/1000 words for the updated guidelines. Comparison of the updated and original text by the translator accounted for this difference. The time for translation by the junior translators was substantially different for the new versus updated guidelines. Because the junior translators were different persons for the new guidelines and the updated guidelines, these comparisons are not meaningful.

Translation of new guidelines by the junior translator and medical proofreader was 2 minutes/1000 words faster with support versus without, but this was not statistically significant (44 minutes/1000 words versus 46 minutes/1000 words; 95% CI -16 to 12; *P*=.94).

For the updated guidelines, the translation by the junior translator and medical proofreader was 59 minutes/1000 words faster with support versus without (152 minutes/1000 words versus 210 minutes/1000 words; 95% CI -115 to -2; *P*=.043 with parametric test and *P*=.053 with nonparametric test).

Translation Quality

The experienced medical translator provided the best quality translations with an HTER score of 3.7 for the new guidelines and 4.2 for the updated guidelines. For the less experienced translators, we did not find any difference in quality in guidelines translated with or without computer assistance. In the test with the new guidelines, the HTER scores were 10.3 with support and 9.5 without. For the updated guidelines, the scores were 5.2 with support and 5.4 without.

On the basis of an analysis of 18 new and updated guidelines (3/arm), we identified 698 revisions by the medical proofreader that related to acceptability and 219 revisions in relation to adequacy. Further details are available in Figures 2 and 3. A third of the revisions were due to terminology. In-depth analysis of the number of terminological revisions per group did not demonstrate any relevant differences.

Discussion

Principal Findings

This study shows that the hybrid approach, that is, human and computer-aided translation, is useful when updates of clinical practice guidelines have to be translated. There was no apparent benefit for the translation of new guidelines. Use of the translation software did not increase the quality of the translations but significantly improved translation speed for updates of existing guidelines. This can be explained by the fact that the translation of unchanged or slightly changed segments is immediately suggested by the translation memory.

Speed of translation is important in the case of updated guidelines. When updating of guidelines is slow, there is an increased risk that guidelines will become out of date, which can affect quality of care. Median times to incorporate new

evidence in updates of guidelines takes 10 months for the Duodecim EBM Guidelines [22]. Translating updated guidelines can increase the use of the guidelines but adds to the delay. It is therefore relevant to find that translation software can contribute to increasing efficiency in this particular case. We did not evaluate the outcomes for guidelines with minor updates, which represent the biggest part of the guideline collection. Our expectation is that here, too, the translation software will enhance efficiency.

The evaluation of adequacy and acceptability demonstrated that the performance of the translation procedures can be improved by the introduction of an automatic terminology consistency check. Terminology was the most important reason for revisions. Contrary to our expectations, the use of the MultiTerm termbases did not result in fewer terminological revisions. An explanation could be that the certificated translator without medical specialization delivered the same overall quality but needed more time to identify the relevant terms. Figure 2 also shows that although the medical proofreader was advised not to change text just for the sake of improving its style, a large number of stylistic changes were made. Providing style guides to the translators and proofreaders might make the process more efficient.

Strengths and Limitations

A limitation of this study is that only 3 translators were involved, with substantial differences in working style. This made it difficult to make meaningful comparisons of the working speed between the different translators. While one of the junior translators worked faster than the experienced medical translator, this was counterbalanced by substantially lower quality scores and more time required for the medical proofread. Another limitation is that there was no monitoring of how the translators used the translation technology. Even though the translators were acquainted with the software, it was not established whether they made optimal use of it. Furthermore, the pilot study used a small guideline sample, and although care was taken to ensure comparability of the texts, this comparability was based on readability scores while arguably other factors may also play a role (eg, some subject matter is more difficult to grasp for the translator than others, regardless of readability scores). In light of these limitations, the results of this study should be interpreted with caution.

Another approach to translating these guidelines would have been to replace online machine translation by Google Translate with a dedicated machine translation component trained on selected bilingual medical data including the validated memory. A limitation of the translation memory approach is that it provides translation support only when there is a sufficient match value between a new sentence and one already stored in the memory. The machine translation approach, on the contrary, is able to combine partial matches into a new translation proposal (but in doing so may also offer more inadequate translation proposals). The Cochrane collaboration is currently using a machine translation approach to translate Cochrane reviews into several languages [23]. Epistemonikos, a multilingual database of the available health evidence, is another project that uses automated statistical machine translations [24]. The performance of our hybrid method supplemented with a dedicated machine translation system remains to be tested.

Because there is only limited evidence on the importance of translation to tackle language barriers, we believe it would be worthwhile to test the effect on reading speed and retention of information. Two previous studies illustrate how this can be tested with the design of a randomized controlled trial [7,8].

Conclusions

The development and updating of guidelines is time consuming and expensive, and strategies are needed to increase cost effectiveness [25]. A large number of clinical practice guidelines and databases with evidence-based point-of-care information are available throughout the world. In a move toward more international collaboration, we expect that the exchange of high-quality guidelines between organizations internationally and the use of translation software can contribute to increasing the cost effectiveness of guidelines. This study provides preliminary evidence to support the usefulness of translation memory technology for keeping a translated set of guidelines up to date. Further research is needed to evaluate the usefulness of dedicated machine translation systems for the translation of new guidelines.

Authors' Contributions

SV, RV, LM, JB, KVO, and KVN conceived and designed the study. LM, JB, and KV analyzed the linguistic data. SV analyzed the time-related data. MG participated as experienced medical translator. JVS participated as medical proofreader. RV, MG, JV, KVN, and BA provided general advice on the study. JB was the guarantor of the study.

Conflicts of Interest

None declared.

References

1. Van de Velde S, Vander SR, Fauquert B, Geens S, Heselmans A, Ramaekers D, et al. EBMPracticeNet: A bilingual national electronic point-of-care project for retrieval of evidence-based clinical guideline information and decision support. JMIR Res Protoc 2013;2(2):e23 [FREE Full text] [doi: 10.2196/resprot.2644] [Medline: 23842038]
2. Sackett D, Rosenberg W, Gray J, Haynes R, Richardson W. Evidence based medicine: What it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-72. [doi: 10.1136/bmj.312.7023.71]

3. Banzi R, Liberati A, Moschetti I, Tagliabue L, Moja L. A review of online evidence-based practice point-of-care information summary providers. J Med Internet Res 2010;12(3):e26 [FREE Full text] [doi: 10.2196/jmir.1288] [Medline: 20610379]
4. Delvaux N, Van de Velde S, Aertgeerts B, Goossens M, Fauquert B, Kunnamo I, et al. Adapting a large database of point of care summarized guidelines: A process description. J Eval Clin Pract 2015. [doi: 10.1111/jep.12426]
5. Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 2013;8:35 [FREE Full text] [doi: 10.1186/1748-5908-8-35] [Medline: 23522377]
6. Blum RH. Pharmaceuticals and Health Policy: International Perspectives on Provision and Control of Medicines. London, UK: Croom Helm; 1981.
7. Gulbrandsen P, Schroeder TV, Milerad J, Nylenna M. Paper or screen, mother tongue or English: Which is better? A randomized trial. JAMA 2002 Jun 5;287(21):2851-2853. [Medline: 12038931]
8. Letelier LM, Zamarin N, Andrade M, Gabrielli L, Caiozzi G, Viviani P, et al. Exploring language barriers to evidence-based health care (EBHC) in post-graduate medical students: A randomised trial. Educ Health (Abingdon) 2007 Nov;20(3):82. [Medline: 18080956]
9. Rosenbaum SE, Glenton C, Cracknell J. User experiences of evidence-based online resources for health professionals: User testing of The Cochrane Library. BMC Med Inform Decis Mak 2008;8:34 [FREE Full text] [doi: 10.1186/1472-6947-8-34] [Medline: 18662382]
10. Eggen R, Tjensvoll K, Nylenna M. Examining the use of an open digital health library for professionals. JMIR Res Protoc 2014;3(4):e66 [FREE Full text] [doi: 10.2196/resprot.3820] [Medline: 25406825]
11. Sheets L, Gavino A, Callaghan F, Fontelo P. Do language fluency and other socioeconomic factors influence the use of PubMed and MedlinePlus? Appl Clin Inform 2013;4(2):170-184 [FREE Full text] [doi: 10.4338/ACI-2013-01-RA-0006] [Medline: 23874356]
12. Desomer A, Dilles T, Steckel S, Duchesnes C, Vanmeerbeek M, Peremans L, et al. Dissemination and Implementation of Clinical Practice Guidelines in Belgium. Brussels, Belgium: KCE; 2013. URL: http://kce.fgov.be/sites/default/files/page_documents/KCE_212_Clinical_practice_guideline_no%20quotes_0_0.pdf [accessed 2015-09-30] [WebCite Cache ID 6bw92eFvE]
13. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [Medline: 18867058]
14. Dallal G. Randomization. 2015. URL: http://www.randomization.com/ [accessed 2015-03-20] [WebCite Cache ID 6XB2M1nw5]
15. Sharon OB. Methodologies for measuring the correlations between post-editing effort and machine translatability. Mach Transl 2005;19(1):37-58.
16. O'Brien S. Machine-Translatability and Post-Editing Effort: An Empirical Study Using Translog and Choice Network Analysis [PhD dissertation]. Dublin, Ireland: Dublin City University; 2006.
17. An Automated Clarity Tool for Dutch and English Discourse. 2012. URL: http://www.lt3.ugent.be/en/projects/hendi/ [accessed 2015-03-20] [WebCite Cache ID 6XB2SBG2q]
18. Time stamp. 2015. URL: http://time-stamp.software.informer.com/ [accessed 2015-03-20] [WebCite Cache ID 6XB2V8qsp]
19. Snover M, Dorr B, Schwartz R, Micciulla L, Makhoul J. A study of translation edit rate with targeted human annotation. College Park, MD: University of Maryland; 2006. URL: https://www.cs.umd.edu/~snover/pub/amta06/ter_amta.pdf [accessed 2015-09-30] [WebCite Cache ID 6bw9jKZvs]
20. Daems J, Macken L, Vandepitte S. Quality as the sum of its parts: A two-step approach for the identification of translation problems and translation quality assessment for ht and mt+pe. In: Proceedings of MT Summit XIV Workshop on Post-Editing Technology and Practice. 2013 Presented at: MT Summit XIV Workshop on Post-editing Technology and Practice; 2013; Nice, France p. 63-71.
21. House J. In: Blum-Kulka S, editor. Interlingual and Intercultural Communication (Tubinger Beitrage Zur Linguistik; No. 272). Tübingen, Germany: John Benjamins Pub Co; 1986.
22. Banzi R, Cinquini M, Liberati A, Moschetti I, Pecoraro V, Tagliabue L, et al. Speed of updating online evidence based point of care summaries: Prospective cohort analysis. BMJ 2011 Sep 23;343(sep22 2):d5856. [doi: 10.1136/bmj.d5856]
23. von Elm E, Ravaud P, Maclehose H, Mbuagbaw L, Garner P, Ried J, et al. Translating cochrane reviews to ensure that healthcare decision-making is informed by high-quality research evidence. PLoS Med 2013;10(9):e1001516.
24. Rada G, Pérez D, Capurro D. Epistemonikos: A free, relational, collaborative, multilingual database of health evidence. Stud Health Technol Inform 2013;192:486-490. [Medline: 23920602]
25. Grol R. Has guideline development gone astray? Yes. BMJ 2010;340:c306. [Medline: 20118175]

Abbreviations

EBM: evidence-based medicine
HTER: Human Translation Edit Rate
Termbase: terminology database

Edited by G Eysenbach; submitted 20.03.15; peer-reviewed by R Banzi, N Mench Bressan; comments to author 02.08.15; revised version received 13.08.15; accepted 18.09.15; published 09.10.15.

Please cite as:
Van de Velde S, Macken L, Vanneste K, Goossens M, Vanschoenbeek J, Aertgeerts B, Vanopstal K, Vander Stichele R, Buysschaert J
Technology for Large-Scale Translation of Clinical Practice Guidelines: A Pilot Study of the Performance of a Hybrid Human and Computer-Assisted Approach
JMIR Med Inform 2015;3(4):e33
URL: http://medinform.jmir.org/2015/4/e33/
doi:10.2196/medinform.4450
PMID:26453372

Short Paper

Resident Use of Text Messaging for Patient Care: Ease of Use or Breach of Privacy?

Micah T Prochaska[1*], MD; Amber-Nicole Bird[2*], MD; Amar Chadaga[3*], MD; Vineet M Arora[4*], MD, MAPP

[1]Section of Hospital Medicine, Department of Medicine, University of Chicago, Chicago, IL, United States

[2]General Internal Medicine, Department of Medicine, Perelman School of Medicine at University of Pennsylvania, Philadelphia, PA, United States

[3]Associate Program Director Internal Medicine Residency Program, Advocate Christ Medical Center, Chicago, IL, United States

[4]Section of General Internal Medicine, Department of Medicine, University of Chicago, Chicago, IL, United States

[*]all authors contributed equally

Corresponding Author:
Micah T Prochaska, MD
Section of Hospital Medicine
Department of Medicine
University of Chicago
5841 South Maryland Avenue
MC 5000
Chicago, IL, 60637
United States
Phone: 1 773 702 6988
Fax: 1 773 759 7398
Email: mprochas@medicine.bsd.uchicago.edu

Abstract

Background: Short message service (SMS) text messaging is an efficient form of communication and pervasive in health care, but may not securely protect patient information. It is unclear if resident providers are aware of the security concerns of SMS text messaging when communicating about patient care.

Objective: We sought to compare residents' preferences for SMS text messaging compared with other forms of in-hospital communication when considering security versus ease of use.

Methods: This study was a cross-sectional multi-institutional survey of internal medicine residents. Residents ranked different communication modalities based on efficiency, ease of use, and security using a Likert scale. Communication options included telephone, email, hospital paging, and SMS text messaging. Respondents also reported whether they had received confidential patient identifiers through any of these modalities.

Results: SMS text messaging was preferred by 71.7% (94/131) of respondents because of its efficiency and by 79.8% (103/129) of respondents because of its ease of use. For security, 82.5% (104/126) of respondents preferred the hospital paging system, whereas only 20.6% (26/126) of respondents preferred SMS text messaging for secure communication. In all, 70.9% (93/131) of respondents reported having received patient identifiers (first and/or last name), 81.7% (107/131) reported receiving patient initials, and 50.4% (66/131) reported receiving a patient's medical record number through SMS text messages.

Conclusions: Residents prefer in-hospital communication through SMS text messaging because of its ease of use and efficiency. Despite security concerns, the majority of residents reported receiving confidential patient information through SMS text messaging. For providers, it is possible that the benefits of improved in-hospital communication with SMS text messaging and the presumed improvement in the coordination and delivery of patient care outweigh security concerns they may have. The tension between the security and convenience of SMS text messaging may represent an educational opportunity to ensure the compliance of mobile technology in the health care setting.

(JMIR Med Inform 2015;3(4):e37) doi:10.2196/medinform.4797

KEYWORDS

in-hospital communication; SMS text messaging; mobile technology

Introduction

Mobile technology (mobile phones and tablets) has been shown to improve physician efficiency [1] and residents perceive it to improve inpatient communication [2-4]. Short message service (SMS) text messaging is one form of communication using mobile technology that is easy to use, accessible, and allows for the rapid and direct transfer of clinical information between providers. Therefore, SMS text messaging has become pervasive in health care [5] and is preferred for in-hospital communication between residents compared to a traditional in-hospital paging system [6]. Yet, SMS text messaging is discouraged by the Joint Commission for Healthcare Communication for security reasons [7] because there are serious concerns about its compliance with the US Health Insurance Portability and Accountability Act (HIPAA) and its ability to protect confidential patient health information when used on personal mobile devices [8].

Currently, it is unclear if the millennial generation of residents, who are comfortable with the ubiquity of SMS text messaging and its benefits, share the preceding concerns regarding SMS text messaging and patient confidentiality. Protecting patient confidentiality is a professional responsibility outlined in the ABIM Foundation physician charter on medical professionalism [9]. Examining residents' understanding of if and how SMS text messaging may violate their obligation to patient confidentiality is one way of evaluating resident professionalism. Additionally, because behaviors learned and developed during medical training are often carried into future practice, it is particularly important to understand residents' perceptions on the use of technology with respect to patient confidentiality [10,11].

Therefore, our study aimed to understand internal medicine residents' preferences for SMS text messaging versus other available in-hospital communication modalities when considering efficiency, ease of use, and security. Additionally, we sought to determine residents' experiences and perceptions of receiving confidential patient information through SMS text messaging.

Methods

A cross-sectional paper survey was administered to internal medicine residents at 2 academic medical centers, one community-based and the other university-based, during the 2013-2014 academic year. Surveys were passed out to individual residents and collected during morning report and noon conference on different days to ensure that all residents willing to participate had the opportunity. The 2 surveyed institutions maintain residency programs represented by equal numbers of males and females. The hospital paging system with telephone call back was the institutionally preferred and supported method of provider communication at both institutions and neither institution supported or endorsed any form of SMS text messaging (including secure text messaging apps). Residents at both institutions were provided institutional emails, iPads (Cupertino, CA) for use in patient care, and at one institution on-call residents were also provided with portable phones for communication. The survey asked residents to rank on a 4-point Likert scale (1 was most preferred and 4 was least preferred) their preferred form of communication when considering efficiency, the ease of use, and the security of the communication modality. Responses were then dichotomized to represent either "preferred" or "not preferred." Communication options included telephone, email, alphanumeric text (hospital) paging system, and SMS text messaging. Respondents were also asked to report whether they had received confidential patient identifiers (name, patient initials, or medical record numbers) through any of the these modalities.

Results

The overall response rate was 76.3% (132/173). For overall efficiency, 71.7% (94/131) of respondents preferred SMS text messaging, whereas 79.8% (103/129) of respondents reported SMS text messaging to be their preferred communication modality with respect to ease of use when communicating with other providers (Figure 1). In comparison, approximately one-third (35.6%, 46/129) of respondents preferred the current hospital paging system for ease of use when communicating with other providers. However, most (82.5%, 104/126) respondents rated the hospital paging system their preferred form of communication for security, whereas only 20.6% (26/126) of respondents preferred SMS text messaging for secure communication. Despite the security concerns of SMS text messaging, 70.9% (93/131) of respondents reported having received protected patient identifiers, including a patient's first and/or last name, through SMS text messages (Figure 2). Many (81.7%, 107/131) reported receiving patient initials through SMS text messages and half (50.4%, 66/131) reported receiving a patient's medical record number through SMS text messages. Responses did not vary by site.

Figure 1. Preference for communication modality comparing ease of use, efficiency, and security.

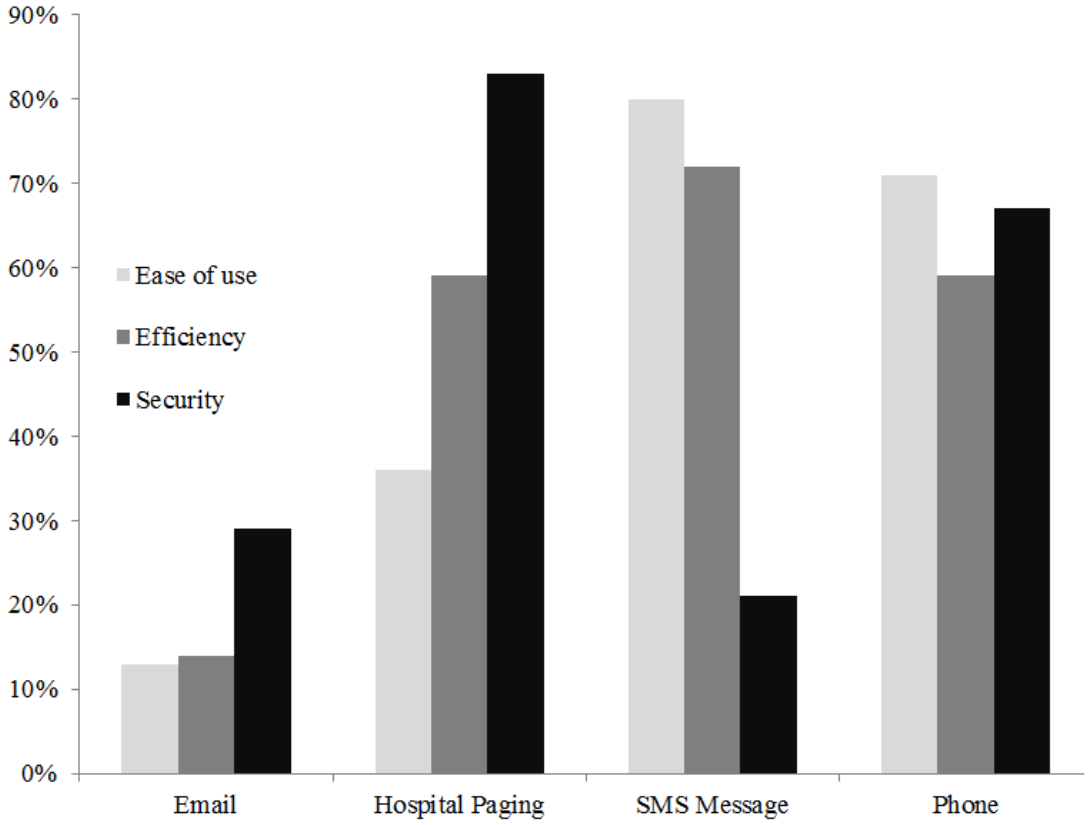
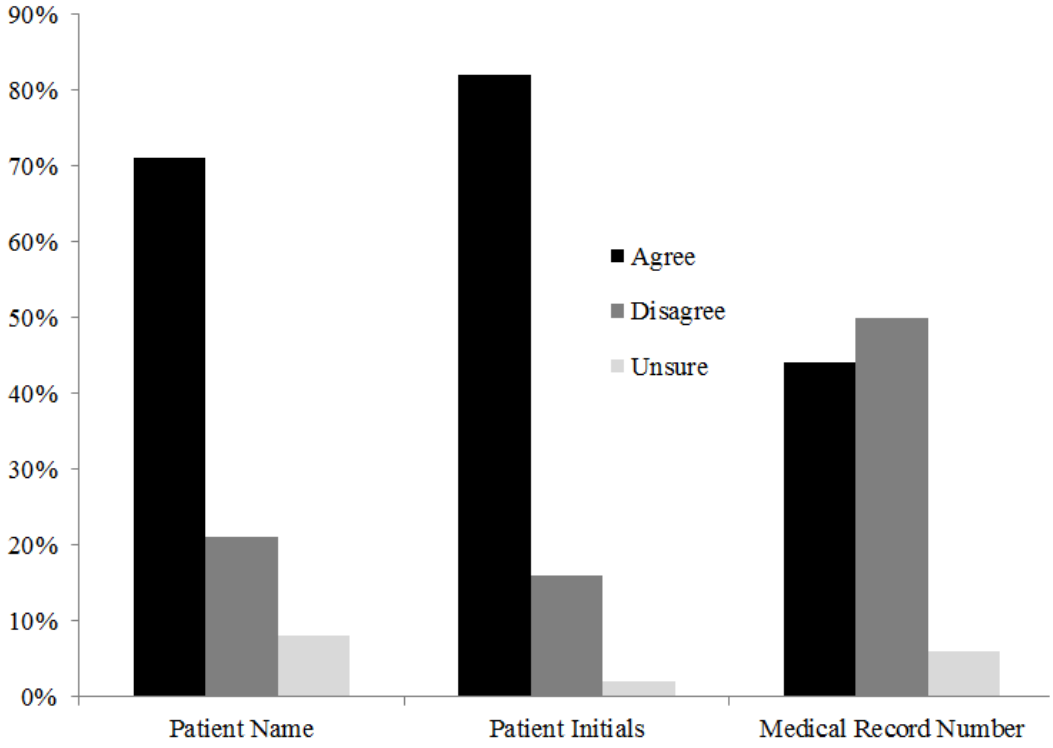


Figure 2. Received protected health information through SMS text messaging.



Discussion

Our data demonstrate that residents are aware of and concerned about the security of SMS text messaging, but prefer it for in-hospital communication because of its efficiency and ease of use. Despite these security concerns, a majority of residents reported receiving confidential patient information through SMS text messaging. One possible explanation for these results is that residents are faced with balancing the tradeoff between the presumed benefits of efficient and easier-to-use modes of

in-hospital communication versus their belief about the security risk posed by communicating protected health information through the different available modes of communication. Interestingly, a majority of residents rated the hospital paging system their preferred method of communication with regards to security, although hospital pagers themselves are not HIPAA compliant [6]. The discrepancy in perception of the security risks of SMS text messaging compared to hospital paging may be due to an underappreciation of the risk of SMS text messaging and an overconfidence in the security of the paging system because it is institutionally supported by the hospital.

However, consequences exist if residents are individually balancing the tradeoff between the benefits of a technology such as SMS text messaging and the security risk it poses to protecting patient information. Residents or trainees may not be accurately estimating either the benefits of SMS text messaging or the real risks and consequences of a health care data breach [12,13]. Additionally, the pressure to be an efficient resident may cause some residents to utilize SMS text messaging in order to maximize efficiency despite the risks to patient confidentiality. In circumstances in which the use of SMS text messaging threatens confidentiality, it also threatens resident professionalism.

Therefore, this presents an educational opportunity to foster understanding about how HIPAA applies to new technologies such as SMS text messaging as well as to inform trainees about the true risks and consequences of data breaches involving protected health information [13]. HIPAA does not specifically ban SMS text messaging or other technologies, but it requires that any exchange of electronic health information meet the minimum standard for physical, network, and process security [14]. By not banning specific technologies, these expectations recognize the fact that new technologies can improve the efficiency and quality of care, but they require that providers and health systems together account for the rights of patients to have their information protected. Therefore, educators have a responsibility to help residents as frontline patient providers and not leave them isolated or at risk with the use of emerging technology. Rather, residents should receive formal education in the standards regarding technology and health care security. Additionally, they should also be engaged in finding and promoting technologies within their institutions, such as secure SMS text messaging apps that are both HIPAA compliant as well as efficient and easy to use. Lastly, residency program directors and institutions should strongly consider understanding the patterns of communication use among residents to ensure that resident practice is in-line with their hospital policy and that hospital policy supports technologies that are efficient, easy to use, and secure for communication between clinicians.

Our study is limited as a 2-institution study and it is possible that our results may not be generalizable to other institutions. Additionally, we collected self-reported data that may be subject to a socially desirable response bias. A socially desirable response bias would make respondents less likely to report having received confidential patient information through SMS text messages, which may mean our data underestimate the true frequency of this phenomenon. Additionally, our survey did not account for the possibility that resident communication preferences may vary based on which member of the medical team they are communicating with and that it is unlikely a resident could SMS text message another member of the medical team with whom they have had no previous contact.

We believe we are the first to study residents' perception of the security of different communication modalities. Our findings suggest that although previous literature supports residents' preference for SMS text messaging, residents are also aware of the security concerns of text messaging. However, the efficiency and ease of use of SMS text messaging when coordinating inpatient care may trump concerns that it does not adequately protect confidential patient information. The tension between the efficiency of a personal technology adapted into health care, but not designed to meets its security standards, will continue to arise as new technologies are developed. As the benefits of these technologies become manifest, we believe it is unrealistic to expect residents or other providers to abstain from their use or self-govern without proper continual education and institutional support to promote awareness of the complexities and nuances of technology and security in health care.

Authors' Contributions

All authors contributed to this work.

Conflicts of Interest

None declared.

References

1. Patel BK, Chapman CG, Luo N, Woodruff JN, Arora VM. Impact of mobile tablet computers on internal medicine resident efficiency. Arch Intern Med 2012 Mar 12;172(5):436-438. [doi: 10.1001/archinternmed.2012.45] [Medline: 22412110]
2. O'Connor C, Friedrich JO, Scales DC, Adhikari NK. The use of wireless e-mail to improve healthcare team communication. J Am Med Inform Assoc 2009;16(5):705-713 [FREE Full text] [doi: 10.1197/jamia.M2299] [Medline: 19567803]
3. Wu RC, Morra D, Quan S, Lai S, Zanjani S, Abrams H, et al. The use of smartphones for clinical communication on internal medicine wards. J Hosp Med 2010;5(9):553-559. [doi: 10.1002/jhm.775] [Medline: 20690190]
4. Wu R, Lo V, Morra D, Appel E, Arany T, Curiale B, et al. A smartphone-enabled communication system to improve hospital communication: usage and perceptions of medical trainees and nurses on general internal medicine wards. J Hosp Med 2015 Feb;10(2):83-89. [doi: 10.1002/jhm.2278] [Medline: 25352429]

5. Frizzell JD, Ahmed B. Text messaging versus paging: new technology for the next generation. J Am Coll Cardiol 2014 Dec 23;64(24):2703-2705. [doi: 10.1016/j.jacc.2014.11.001] [Medline: 25524347]
6. Przybylo JA, Wang A, Loftus P, Evans KH, Chu I, Shieh L. Smarter hospital communication: secure smartphone text messaging improves provider satisfaction and perception of efficacy, workflow. J Hosp Med 2014 Sep;9(9):573-578 [FREE Full text] [doi: 10.1002/jhm.2228] [Medline: 25110991]
7. Pelletier M. The Joint Commission. 2015 Apr 10. To text or not to text? URL: http://www.jointcommission.org/the_view_from_the_joint_commission/to_text_or_not_to_text/default.aspx [accessed 2015-06-02] [WebCite Cache ID 6YzW0ujzj]
8. Boulton C. The Wall Street Journal. 2014 Sep 05. Hospitals have weak security for mobile devices: Forrester URL: http://blogs.wsj.com/cio/2014/09/05/hospitals-have-weak-security-for-mobile-devices-forrester/ [WebCite Cache ID 6YzVlMzlm]
9. American Board of Internal Medicine Foundation. Professionalism Charter URL: http://www.abimfoundation.org/Professionalism/Physician-Charter/Commitments-of-the-Charter.aspx [accessed 2015-06-02] [WebCite Cache ID 6YzVfISEg]
10. Teherani A, Hodgson CS, Banach M, Papadakis MA. Domains of unprofessional behavior during medical school associated with future disciplinary action by a state medical board. Acad Med 2005 Oct;80(10 Suppl):S17-S20. [Medline: 16199450]
11. Papadakis MA, Arnold G, Blank L, Holmboe E, Lipner R. Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards. Ann Intern Med 2008 Jun 3;148(11):869-876. [Medline: 18519932]
12. Abelson R, Creswell J. The New York Times. 2015 Feb 06. Data breach at Anthem may forecast a trend URL: http://www.nytimes.com/2015/02/07/business/data-breach-at-anthem-may-lead-to-others.html?_r=0 [WebCite Cache ID 6YzVi2N7w]
13. Tran K, Morra D, Lo V, Quan SD, Abrams H, Wu RC. Medical students and personal smartphones in the clinical environment: the impact on confidentiality of personal health information and professionalism. J Med Internet Res 2014;16(5):e132 [FREE Full text] [doi: 10.2196/jmir.3138] [Medline: 24855046]
14. US Department of Health & Human Services. Summary of the HIPAA security rule URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/srsummary.html [WebCite Cache ID 6YzWNkdE9]
15. HHS.gov. Summary of the HIPAA Security Rule URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/srsummary.html [WebCite Cache ID 6YzWNkdE9]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
SMS: short message service

Edited by G Eysenbach; submitted 03.06.15; peer-reviewed by L Shieh, R Wu; comments to author 16.08.15; accepted 23.09.15; published 26.11.15.

Please cite as:
Prochaska MT, Bird AN, Chadaga A, Arora VM
Resident Use of Text Messaging for Patient Care: Ease of Use or Breach of Privacy?
JMIR Med Inform 2015;3(4):e37
URL: http://medinform.jmir.org/2015/4/e37/
doi:10.2196/medinform.4797
PMID:26611620

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org