CONTENTS

**Original Papers**

The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature (e1)
Suhila Sawesi, Mohamed Rashrash, Kanitha Phalakornkule, Janet Carpenter, Josette Jones. ........................................ 3

Early Indication of Decompensated Heart Failure in Patients on Home-Telemonitoring: A Comparison of Prediction Algorithms Based on Daily Weight and Noninvasive Transthoracic Bio-impedance (e3)
Illapha Cuba Gyllensten, Alberto Bonomi, Kevin Goode, Harald Reiter, Joerg Habetha, Oliver Amft, John Cleland. .................................................................................................................. 27

Health Information Technology: Meaningful Use and Next Steps to Improving Electronic Facilitation of Medication Adherence (e9)
Hayden Bosworth, Leah Zullig, Phil Mendys, Michael Ho, Troy Trygstad, Christopher Granger, Megan Oakes, Bradi Granger. ........................................................................................................ 40

Conceptual Models in Health Informatics Research: A Literature Review and Suggestions for Development (e7)
Kathleen Gray, Paulina Sockolow. ............................................................................................................................. 49

Integration of Provider, Pharmacy, and Patient-Reported Data to Improve Medication Adherence for Type 2 Diabetes: A Controlled Before-After Pilot Study (e4)
Brian Dixon, Abdullah Alzeer, Erin Phillips, David Marrero. ....................................................................................... 58

Industry and Occupation in the Electronic Health Record: An Investigation of the National Institute for Occupational Safety and Health Industry and Occupation Computerized Coding System (e5)
Matthew Schmitz, Linda Forst. ................................................................................................................................. 71

Disruptive Innovation: Implementation of Electronic Consultations in a Veterans Affairs Health Care System (e6)
Gouri Gupte, Varsha Vimalananda, Steven Simon, Katerina DeVito, Justice Clark, Jay Orlander. ...................... 80

Improving Inpatient Surveys: Web-Based Computer Adaptive Testing Accessed via Mobile Phone QR Codes (e8)
Tsair-Wei Chien, Weir-Sen Lin. ................................................................................................................................. 90
Quantification of Subcutaneous and Visceral Adipose Tissue From Computed Tomography Scans: Development and Validation Study (22)
Young Kim, Ji Park, Jong Kim, Chan-Soo Park, John Gonzalez, Seung Lee, Kwang Kim, Jae Oh. ............... 100
The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature

Suhila Sawesi¹, B Pharm, MPharmSc, PhD (c); Mohamed Rashrash², B Pharm, MPharm, PhD; Kanitha Phalakornkule¹, MSc, PhD (c); Janet S Carpenter³, RN, PhD; Josette F Jones³, RN, PhD

¹School of Informatics and Computing – Indianapolis, Department of BioHealth Informatics, IUPUI, Indianapolis, IN, United States
²School of Pharmacy, Chapman University, Irvine, CA, United States
³Indiana University School of Nursing, Science of Nursing Care Department, Indiana University, Indianapolis, IN, United States

Corresponding Author:
Suhila Sawesi, B Pharm, MPharmSc, PhD (c)
School of Informatics and Computing – Indianapolis
Department of BioHealth Informatics
IUPUI
Walker Plaza (WK) 120
719 Indiana Avenue
Indianapolis, IN, 46202
United States
Phone: 1 317 274 8059
Fax: 1 317 274 8059
Email: ssawesi@umail.iu.edu

Abstract

Background: Advancements in information technology (IT) and its increasingly ubiquitous nature expand the ability to engage patients in the health care process and motivate health behavior change.

Objective: Our aim was to systematically review the (1) impact of IT platforms used to promote patients’ engagement and to effect change in health behaviors and health outcomes, (2) behavior theories or models applied as bases for developing these interventions and their impact on health outcomes, (3) different ways of measuring health outcomes, (4) usability, feasibility, and acceptability of these technologies among patients, and (5) challenges and research directions for implementing IT platforms to meaningfully impact patient engagement and health outcomes.

Methods: PubMed, Web of Science, PsycINFO, and Google Scholar were searched for studies published from 2000 to December 2014. Two reviewers assessed the quality of the included papers, and potentially relevant studies were retrieved and assessed for eligibility based on predetermined inclusion criteria.

Results: A total of 170 articles met the inclusion criteria and were reviewed in detail. Overall, 88.8% (151/170) of studies showed positive impact on patient behavior and 82.9% (141/170) reported high levels of improvement in patient engagement. Only 47.1% (80/170) referenced specific behavior theories and only 33.5% (57/170) assessed the usability of IT platforms. The majority of studies used indirect ways to measure health outcomes (65.9%, 112/170).

Conclusions: In general, the review has shown that IT platforms can enhance patient engagement and improve health outcomes. Few studies addressed usability of these interventions, and the reason for not using specific behavior theories remains unclear. Further research is needed to clarify these important questions. In addition, an assessment of these types of interventions should be conducted based on a common framework using a large variety of measurements; these measurements should include those related to motivation for health behavior change, long-standing adherence, expenditure, satisfaction, and health outcomes.

(JMIR Med Inform 2016;4(1):e1) doi:10.2196/medinform.4514

Keywords
patient engagement; patient behavior; technology; Internet; web-based; cell phone; social media
Introduction

Patient engagement is currently considered the cornerstone of the health care system revolution for its positive impact on health outcomes and health care costs [1,2]. A growing body of evidence demonstrates that lack of patient engagement is a major contributor to preventable deaths. In fact, it is estimated that 40% of deaths in the United States are caused by modifiable behavioral issues, including smoking, obesity, poor blood sugar control, poor blood pressure control, inadequate exercise, medication non-adherence, and neglect in attending follow-up medical appointments [3]. As a result, patients must be encouraged to become more involved with managing their own care. Frequent, real-time communication and feedback are essential in supporting health behavior change and empowering patient engagement in the health care process [4]. However, the traditional model of care delivery, a face-to-face interaction with an expert or trusted health care provider, can be implemented only with a small number of patients and thus has limited impact and limited reach [5]. In an effort to reach and engage larger numbers of patients, researchers and clinicians have begun exploring the role of information technology (IT) platforms in patient engagement and health behavior change interventions [5,6]. It is assumed that face-to-face interaction in the traditional model can be mimicked by peer-to-peer or peer group support in social media.

IT platforms are being embraced as a way to enhance patient engagement in the health care process, improve quality of care, support health care safety, and provide cost-effective health services for patients [6-9]. Numerous IT platforms are used to motivate patient engagement in health behavior change including short message service (SMS)-capable mobile devices, Internet-based interventions, social media, and other online communication tools [10-12]. Previous systematic reviews have evaluated the potential benefit of IT platforms in managing different health conditions and how these platforms have been used to actively engage patients and change unhealthy patient behavior. A systematic review conducted to assess the effectiveness of IT platforms on physical activity and dietary behavior change found that 51% of studies showed positive results, although a significant proportion of the studies showed no significant effect [13]. The reviewed interventions tended to focus on specific technology (eg, desktop applications), while mobile devices, such as mobile phones and text messaging devices were not included. Similarly, Webb et al reviewed 85 studies on the impact of Internet-based interventions on health behavior change and found small but significant effects on health-related behavior, especially with regards to interventions grounded in behavioral theory. Although the review mentioned that the effectiveness of Internet-based interventions was enhanced by using additional IT methods, such as text messaging (SMS), it did not focus on the distinction between these different interventions [8].

In addition, a meta-analysis performed to investigate the effectiveness of Web-based interventions on health behavior changes found that Web-based interventions improve patient outcomes. This particular meta-analysis, however, referred only to Web-based interventions in specific problem areas and focused on a relatively narrow range of technologies [14]. A recent systematic review that investigated the effectiveness of the IT platform on self-management among diabetic patients showed positive effects in 74% of studies [15]. Another research study showed that successful health behavior interventions may contribute to understanding of health behavior theories and their appropriate use [16]. Mobile-based interventions and Web-based interventions developed based on health behavior theories are more likely to effectively change patient health behavior and maintain behavior change than non-theory-based interventions [8,17-19]. Basing IT interventions on behavior theories can help test and detect why interventions succeed or fail [20]. Health behavioral theories can identify key determinants of the target behaviors and identify behavior change strategies essential to obtain desired health outcomes; this knowledge can then be transformed into specific behavioral strategies that patients can adapt in their daily life [20].

Conclusions drawn from these reviews are important; they provide insights but no clear answers about the effectiveness of IT platforms on patient engagement and behavior change. They do not address which interventions are used most or are most effective with which theory or model when it comes to improving patients’ health behaviors and patient engagement. IT platforms generally can have high potential benefits and some proven effects; however, specific components in several health conditions associated with success remain unclear. To better understand how to build a successful intervention that can engage patients to change their behavior meaningfully, we performed a systematic review.

Review aims were to systematically determine (1) the impact of IT platforms used to promote patient engagement and to effect change in health behaviors and health outcomes, (2) behavioral theories or models applied as bases for developing these interventions and their impact on health outcomes, (3) different ways of measuring health outcomes, (4) usability, feasibility, and acceptability of these technologies among patients, and (5) challenges and research directions for implementing IT platforms to meaningfully impact patient engagement and health outcomes.

Methods

Search Strategy and Data Sources

Electronic literature searches were performed using four databases: PubMed, Web of Science, PsycINFO, and Google Scholar. Google Scholar was searched because it had sufficiently wide coverage to be used instead of several databases [21-23]. The reference lists of retrieved articles from searches were screened for additional articles. Searches used the following medical subject headings (MeSH) terms in various combinations: patient engagement, health, promotion, behavior, digital, technology, email, Internet, Web-based, cell phone, social media, computer, and intervention.

Inclusion and Excluding Criteria

The following criteria were used to select the articles: (1) all types of study designs published in scientific journals between 2000 and December 2014 were included, excluding conference
proceedings, book chapters, reviews, dissertations, and protocols. (2) studies that evaluated and reported the impact of health information technology platforms on patients’ health outcome, (3) studies that focused on disease management rather than more general health promotion including but not limited to patient education, symptom monitoring, medication adherence, diet, and physical activity, (4) studies that addressed patient engagement and health-related behavior change through the use of IT platforms such as social networking sites, mobile telephony, video and teleconferencing, email, SMS, and electronic monitoring, (5) studies that explored different factors affecting patient engagement and health behavior change were excluded, (6) studies that were published in languages other than English were excluded, (7) studies where the patient was not the main actor (i.e., studies that were clinician-focused), and (8) the methodological quality (see Multimedia Appendix 1) of articles was evaluated to establish their inclusion in the review using 10 items adopted from Critical Appraisal Skills Programme (CASP) [24,25]. The criteria that were used in the quality assessment included (1) study name, (2) aims clearly stated, (3) appropriate research design, (4) appropriate recruitment strategy, (5) theories clearly stated, (6) usability tested within the study, (7) patient engagement part of study, (8) appropriate data collection method, (9) data analysis sufficiently rigorous, and (10) findings clearly stated. After the completion of the methodological quality assessment, the studies that met the criteria for the categories of “good” were reviewed (i.e., bad=0-33%, satisfactory=34-66%, and good=67-100%) [26].

Data Extraction

Two investigators independently reviewed the titles and then abstracts. The same investigators read and screened for full text eligibility. Data extraction was carried out by 1 reviewer and was rechecked for accuracy by another reviewer. The reasons for exclusion were recorded. Discrepancies were resolved by joint probability of agreement (0.98) [15].

A meta-analysis was not feasible due to the varying data collection methods and outcome measures. Therefore, eligible studies were broken down and evaluated in a narrative format using some statistical analysis when feasible and summarized systematically according to the following key information abstracted from them: study details (including author name, year, country, and study design); study characteristics (including sample size and condition/disease); intervention details (including technology used and duration); and outcome details (including direct and indirect assessment methods); and impact of intervention, usability assessment, patient engagement, and theory used in interventions classified according to Leventhal (biomedical model, behavioral learning, communicative, cognitive theory, and self-regulative) [27-29].

The outcomes variable was classified into (1) positive impact in which health information technology platform was associated with improvement in one or more aspects of care and (2) no impact or no noticeable improvement or change in health outcomes. This was assessed based on the overall conclusion made by the authors of each study. Most studies used statistical methods to test hypotheses or describe quantitative findings.

Patient engagement was measured based on the overall conclusion. This was usually measured by timed patient log-ins, communication with the health care provider via secure message, or data download.

Results

Search and Selection Results

Figure 1 shows the flow chart that describes the process of identifying the relevant literature. A separate comprehensive search using 4 databases yielded 2235 articles. Following removal of duplicates, our search identified 786 potentially relevant articles. These were scanned keeping 219 papers for full reading at full text level, of which 59 were screened and rejected, leaving 160 studies to be included in the review. Ten additional papers were included from the reference lists of retrieved articles. A total of 170 articles matched the initial search criteria.
Figure 1. Flow diagram of included and excluded studies.

### Article Characteristics

Table 1 provides a summary of the studies, and Multimedia Appendix 2 summarizes the 170 articles included in the research study and their characteristics. Multimedia Appendix 1 summarizes the quality assessment of the 170 included articles. Multimedia Appendix 3 contains definitions of terms used in the paper. Different categories of IT platforms were identified including Internet-based interventions (50.6%, 86/170), mobile-based interventions (25.9%, 44/170), social media (9.4%, 16/170), video game technology (3.5%, 6/170), and telemonitoring (10.6%, 18/170). Publication years ranged from 2000 to 2014, with an overall increase in articles published more recently (21.8%, 37/170 in 2014). The majority of studies were implemented in the United States (54.7%, 93/170). With respect to the different targeted disorders, hormonal disorders were most frequently targeted (22.4%, 38/170 studies, eg, diabetes). The literature was dominated by randomized controlled trials (65.9%, 112/170). The duration of these studies ranged from 1 week to 48 months, and sample sizes ranged from 1-22,337 subjects. Articles included in this review were categorized in five topics based on study aims: impact of IT platform on health outcomes, patient engagement in health behavior change, theory of health behavior, ways to assess health outcomes, and usability assessment (see Table 1).
Table 1. Summary of the review results based on types of IT platforms.

<table>
<thead>
<tr>
<th>Health condition, n (%)</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bone, joint, and muscle disorders</td>
<td>3 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain, spinal cord, and nerve disorders</td>
<td>7 (8)</td>
<td>1 (2)</td>
<td>2 (33)</td>
<td>1 (6)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Cancer</td>
<td>5 (8)</td>
<td>2 (5)</td>
<td>1 (17)</td>
<td>2 (13)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Disorders of nutrition and metabolism</td>
<td>13 (15)</td>
<td>4 (9)</td>
<td>1 (17)</td>
<td>2 (13)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Ears, nose, and throat disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health hazard</td>
<td>5 (6)</td>
<td>6 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart and blood vessel disorders</td>
<td>5 (6)</td>
<td>3 (7)</td>
<td></td>
<td></td>
<td>6 (33)</td>
</tr>
<tr>
<td>Hormonal disorders</td>
<td>20 (23)</td>
<td>11 (25)</td>
<td>4 (25)</td>
<td></td>
<td>3 (17)</td>
</tr>
<tr>
<td>Immune disorders</td>
<td>4 (5)</td>
<td>5 (11)</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Lung and airway disorders</td>
<td>2 (2)</td>
<td>1 (2)</td>
<td>1 (17)</td>
<td>1 (6)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Mental health disorders</td>
<td>12 (14)</td>
<td>4 (9)</td>
<td>1 (17)</td>
<td>2 (13)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Skin disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women’s health issues</td>
<td>3 (3)</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not specified</td>
<td>7 (8)</td>
<td>4 (9)</td>
<td></td>
<td>2 (13)</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Country, n (%)                               |                 |              |                  |                       |                        |
| Australia                                    | 7 (8)           | 5 (42)       |                  |                       |                        |
| Austria                                      |                 |              |                  |                       | 1 (6)                  |
| Bangladesh                                   |                 |              |                  |                       |                        |
| Canada                                       | 4 (5)           |              |                  |                       | 2 (11)                 |
| Chile                                        | 1 (1)           |              |                  |                       |                        |
| China                                        |                 |              |                  |                       |                        |
| France                                       |                 |              |                  |                       |                        |
| Germany                                      | 3 (3)           |              |                  |                       |                        |
| Israel                                       |                 |              |                  |                       | 1 (6)                  |
| Italy                                        | 1 (1)           |              | 1 (2)            |                       |                        |
| Japan                                        | 1 (1)           |              |                  | 1 (6)                 |                        |
| Kenya                                        |                 |              |                  |                       |                        |
| Korea                                        | 1 (1)           |              | 1 (2)            |                       | 1 (6)                  |
| Malaysia                                     |                 |              |                  |                       |                        |
| Netherlands                                  | 4 (5)           |              | 1 (17)           |                       | 2 (11)                 |
| New Zealand                                  |                 |              |                  |                       |                        |
| Norway                                       | 1 (2)           |              |                  |                       |                        |
| Poland                                       |                 |              |                  |                       | 1 (6)                  |</p>
<table>
<thead>
<tr>
<th>Country</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russia</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td>1 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Korea</td>
<td>2 (2)</td>
<td>4 (5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td>1 (6)</td>
</tr>
<tr>
<td>Sweden</td>
<td>2 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>1 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>5 (6)</td>
<td>7 (16)</td>
<td>1 (17)</td>
<td>1 (6)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>53 (62)</td>
<td>14 (32)</td>
<td>4 (67)</td>
<td>14 (88)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>United States</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victoria</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Study design, n (%)**

<table>
<thead>
<tr>
<th>Study design</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized controlled trial</td>
<td>55 (64)</td>
<td>34 (30)</td>
<td>2 (33)</td>
<td>7 (44)</td>
<td>14 (78)</td>
</tr>
<tr>
<td>Case study</td>
<td>2 (2)</td>
<td>1 (2)</td>
<td>2 (33)</td>
<td>2 (13)</td>
<td></td>
</tr>
<tr>
<td>Cohort study</td>
<td>10 (12)</td>
<td>4 (5)</td>
<td>1 (17)</td>
<td>1 (6)</td>
<td>3 (17)</td>
</tr>
<tr>
<td>Cross-sectional analysis</td>
<td>8 (9)</td>
<td>1 (2)</td>
<td>5 (31)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Quasi-experimental trial</td>
<td>11 (13)</td>
<td>4 (5)</td>
<td>1 (17)</td>
<td>1 (6)</td>
<td></td>
</tr>
</tbody>
</table>

**Ways to measure health outcomes, n (%)**

<table>
<thead>
<tr>
<th>Ways to measure health outcomes</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>28 (33)</td>
<td>20 (45)</td>
<td>3 (50)</td>
<td>1 (6)</td>
<td>6 (33)</td>
</tr>
<tr>
<td>Indirect</td>
<td>58 (67)</td>
<td>24 (55)</td>
<td>3 (50)</td>
<td>15 (94)</td>
<td>12 (67)</td>
</tr>
</tbody>
</table>

**Impact of technology, n (%)**

<table>
<thead>
<tr>
<th>Impact of technology</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>75 (87)</td>
<td>41 (93)</td>
<td>6 (100)</td>
<td>13 (81)</td>
<td>16 (89)</td>
</tr>
<tr>
<td>No</td>
<td>11 (13)</td>
<td>3 (7)</td>
<td>3 (19)</td>
<td>3 (19)</td>
<td>2 (11)</td>
</tr>
</tbody>
</table>

**Usability assessment, n (%)**

<table>
<thead>
<tr>
<th>Usability assessment</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>38 (44)</td>
<td>8 (18)</td>
<td>1 (17)</td>
<td>8 (50)</td>
<td>3 (17)</td>
</tr>
<tr>
<td>No</td>
<td>48 (56)</td>
<td>36 (82)</td>
<td>5 (83)</td>
<td>8 (50)</td>
<td>15 (83)</td>
</tr>
</tbody>
</table>

**Patient engagement, n (%)**

<table>
<thead>
<tr>
<th>Patient engagement</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>68 (79)</td>
<td>38 (86)</td>
<td>6 (100)</td>
<td>13 (81)</td>
<td>16 (89)</td>
</tr>
<tr>
<td>No</td>
<td>18 (21)</td>
<td>6 (14)</td>
<td>3 (19)</td>
<td>3 (19)</td>
<td>2 (11)</td>
</tr>
</tbody>
</table>

**Theory of behavior change, n (%)**

<table>
<thead>
<tr>
<th>Theory of behavior change</th>
<th>Internet (N=86)</th>
<th>Phone (N=44)</th>
<th>Video game (N=6)</th>
<th>Social network (N=16)</th>
<th>Tele-monitoring (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomedical theory (chronic model)</td>
<td>1 (1)</td>
<td></td>
<td></td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Behavioral learning theory</td>
<td>3 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication (social support theory)</td>
<td>5 (6)</td>
<td></td>
<td>5 (11)</td>
<td>2 (13)</td>
<td></td>
</tr>
<tr>
<td>Cognitive theorya (TPB, SOC, TTM, self-efficacy, information motivation, and behavioral skill)</td>
<td>40 (47)</td>
<td></td>
<td>9 (20)</td>
<td>2 (33)</td>
<td>2 (13)</td>
</tr>
</tbody>
</table>
Impact of IT Platforms on Health Outcomes

Overall, IT platforms have been shown to improve health behavior among different disease categories (88.8%, 151/170), although the majority of the positive impact has been shown among hormonal disorders (20.6%, 35/170) (see Table 2). Among studies utilizing Internet-based platforms, 87% (75/86) of studies showed a significant impact on health outcomes. Studies also showed that the use of Internet-based tailored weight control programs was correlated with significant increases in weight loss [30,31] and walking distance ($P<.05$) [32]. Similarly, mobile-based platforms showed significant effects on health outcomes (91%, 40/44). For example, a study examined use of text messages among patients with diabetes and found a significant decrease in HbA1C level, improved medication adherence, and decreased in emergency service use [33]. Social media showed a positive impact on health outcomes (81%, 13/16). For example, one study indicated that Twitter usage among cancer patients was a valuable medium for sharing information, discussing treatments, and also acted as a psychological support [34]. The use of Facebook has also been found to help improve asthma care [35]. As such, this review found that 100% (6/6) of studies had a positive impact on patient health behavior when implementing a video game as an intervention to change health behavior. In a specific example, King et al concluded that video games can be implemented successfully among hyperfunctional voice disorder as a “voice therapeutic protocol”, a voice and speech therapy program including a set of vocal tasks using syllable repetitions and chanting of songs and phrases [36]. Furthermore, the literature showed that telemonitoring improved health outcomes (89%, 16/18). One telemonitoring-based study assessed the effects of a glucose monitoring system on HbA1c levels in diabetic patients and found that usage of this system was correlated with a significant decrease in HbA1c ($P=.001$) [37]. Another study evaluated the impact of home-based telemonitoring on patients with heart failure and showed a significant correlative improvement in patients’ health outcomes [38,39].

In contrast, 11% of studies (19/170) showed no impact of using IT platforms on health behavior. Among studies using Internet-based platforms, 13% (11/86) did not find significant results. One study using a Web-based behavior change program found no differences in smoking abstinence rates at 3- and 6-month follow-up assessment [40] and no maintenance of weight loss in an Internet-based intervention group compared to the study’s control group [41]. Also, 7% of (3/44) mobile phone studies reported non-significant impact [33,42-44]. Two mobile phone platform studies did not find a significant reduction in HbA1c level among diabetic patients when SMS text messaging was used to manage their health care ($P<.10$) [33,34]. Moreover, 18% (3/16) of studies showed undesirable effects from using social media [35,45,46]. For instance, Kaplan et al found that psychiatric patients who participated in Internet peer support reported higher levels of distress compared to those who did not participate [45]. The literature shows that 12% (2/18) of telemonitoring studies had no effect on health outcomes. One particular study found significant changes in neither readmission rate [47] nor medication adherence [48] among patients with heart failure.
Table 2. Impact of IT platforms among different disorders (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Disorders</th>
<th>Internet</th>
<th>Mobile</th>
<th>Social media</th>
<th>Tele-monitoring</th>
<th>Video game</th>
<th>Total yes</th>
<th>Total no</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bone, joint, and muscle</td>
<td>3 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 (2)</td>
<td>3 (2)</td>
<td></td>
</tr>
<tr>
<td>Brain, spinal cord, and nerves</td>
<td>7 (8)</td>
<td>1 (2)</td>
<td>1 (6)</td>
<td>2 (33)</td>
<td></td>
<td>11 (6)</td>
<td>1 (1)</td>
<td>12 (7)</td>
</tr>
<tr>
<td>Cancer</td>
<td>5 (6)</td>
<td>2 (5)</td>
<td>2 (13)</td>
<td>2 (11)</td>
<td>1 (17)</td>
<td>12 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrition and metabolism</td>
<td>10 (12)</td>
<td>3 (3)</td>
<td>2 (13)</td>
<td>1 (6)</td>
<td>1 (17)</td>
<td>18 (11)</td>
<td>3 (3)</td>
<td>21 (12)</td>
</tr>
<tr>
<td>Ears, nose, and throat</td>
<td></td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td>1 (1)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Eye</td>
<td></td>
<td></td>
<td>1 (6)</td>
<td></td>
<td></td>
<td></td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Health hazard</td>
<td>4 (5)</td>
<td>1 (1)</td>
<td>6 (14)</td>
<td></td>
<td></td>
<td>10 (6)</td>
<td>1 (1)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>Heart and blood vessel</td>
<td>4 (5)</td>
<td>1 (1)</td>
<td>3 (7)</td>
<td>4 (22)</td>
<td>2 (11)</td>
<td>11 (6)</td>
<td>3 (2)</td>
<td>14 (8)</td>
</tr>
<tr>
<td>Hormonal</td>
<td>19 (22)</td>
<td>1 (1)</td>
<td>9 (20)</td>
<td>4 (25)</td>
<td>3 (17)</td>
<td>35 (21)</td>
<td>3 (2)</td>
<td>38 (22)</td>
</tr>
<tr>
<td>Immune system</td>
<td>2 (2)</td>
<td>2 (2)</td>
<td>5 (11)</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td>8 (5)</td>
<td>3 (2)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>Lung and airway</td>
<td>2 (2)</td>
<td>1 (2)</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td>1 (17)</td>
<td>6 (4)</td>
<td></td>
<td>6 (4)</td>
</tr>
<tr>
<td>Mental health</td>
<td>11 (13)</td>
<td>1 (1)</td>
<td>4 (9)</td>
<td>2 (13)</td>
<td>2 (11)</td>
<td>20 (12)</td>
<td>1 (1)</td>
<td>21 (12)</td>
</tr>
<tr>
<td>Not specified</td>
<td>5 (6)</td>
<td>2 (2)</td>
<td>4 (9)</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td>10 (6)</td>
<td>3 (2)</td>
<td>13 (8)</td>
</tr>
<tr>
<td>Skin</td>
<td>1 (2)</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td>2 (1)</td>
<td></td>
<td>2 (1)</td>
</tr>
<tr>
<td>Women’s health</td>
<td>3 (3)</td>
<td>1 (2)</td>
<td></td>
<td></td>
<td></td>
<td>4 (2)</td>
<td></td>
<td>4 (2)</td>
</tr>
<tr>
<td>Total</td>
<td>75 (87)</td>
<td>11 (13)</td>
<td>41 (93)</td>
<td>3 (7)</td>
<td>13 (81)</td>
<td>16 (89)</td>
<td>2 (11)</td>
<td>6 (100)</td>
</tr>
</tbody>
</table>

Patient Engagement

In total, 82.9% (141/170) of studies reported improvement in patient engagement after using IT platforms (see Table 3). Among Internet-based interventions, 79% (68/86) of studies reported a high level of patient engagement. For example, a research study reported that human immunodeficiency virus patients used the Internet-based intervention a majority of the time to access information and manage their health [49,50]. Among studies using mobile-based interventions, 86% (38/44) reported improvement in patient engagement. One mobile-based intervention study found that text messaging enhanced successful engagement of diabetic patients in their own health care. Patients were able to use this study’s text message system for clinical data queries and communicating with health care providers [51]. Similarly, 81% (13/16) of studies reported that social media was helpful in improving patient engagement. One study found that Facebook provided a forum for reporting personal experiences, asking questions, and receiving direct feedback for people living with diabetes [46]. Another study showed that social media was helpful to individuals with lower patient activation [52-54]. In addition, it was found that video games could enhance patients’ active participation in the health care process (100%, 6/6). One video game-based study demonstrated that a health-based video game could help build an effective client-therapist relationship, help structure sessions, and improve patient engagement in the therapeutic process [55,56]. Likewise, the literature showed that telemonitoring has been particularly useful for improving patient engagement remotely (88.8%, 16/18) [57-63], as traditional point-of-care-based ways to monitor patients are costly and difficult to implement [64].

Overall, analysis showed significant correlations between patient engagement in health care and the impact of IT platforms ($\chi^2 = 39.8836, P < .001$). Only Internet-based platforms had a significant association between patient engagement and impact of technology on outcomes ($\chi^2 = 28.2558, P < .001$).
Table 3. Impact of IT platforms on patient engagement (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Impact of IT platforms, n (%)</th>
<th>Total yes</th>
<th>Total no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td>135 (63)</td>
<td>6 (32)</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>16 (37)</td>
<td>13 (68)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>151 (100)</td>
<td>19 (100)</td>
</tr>
</tbody>
</table>

Table 4. Impact of IT platforms and theories of health behavior (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Behavior theory</th>
<th>Impact of IT platforms, n (%)</th>
<th>Total yes</th>
<th>Total no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomedical theory (chronic model)</td>
<td>1 (1)</td>
<td>2 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Behavioral learning theory</td>
<td>2 (2)</td>
<td>2 (1)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Communication (social support theory)</td>
<td>4 (5)</td>
<td>4 (6)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>Cognitive theory (TPB, SOC, TTM, self-efficacy, information motivation, and behavioral skill)</td>
<td>36 (42)</td>
<td>53 (31)</td>
<td>58 (34)</td>
</tr>
<tr>
<td>Self-regulatory</td>
<td>5 (6)</td>
<td>1 (1)</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Total of used theory</td>
<td>48 (56)</td>
<td>71 (42)</td>
<td>80 (47)</td>
</tr>
<tr>
<td>Theory not reported</td>
<td>27 (31)</td>
<td>20 (29)</td>
<td>43 (53)</td>
</tr>
<tr>
<td>Total</td>
<td>75 (87)</td>
<td>151 (89)</td>
<td>170 (100)</td>
</tr>
</tbody>
</table>

Behavior Theory

Overall results showed that 47.0% (80/170) of the literature explicitly referenced theory (see Table 4). Among Internet-based interventions, 64% (55/86) of studies mentioned the use of behavior theories. Cognitive theories dominated this category (47%, 40/86). Further, 32% (14/44) of mobile-based intervention studies reported use of behavior theories. Cognitive theories were also the most widely used among this category (30%, 13/44) [42,51,65-71]. Moreover, 38% (6/16) of social media studies used behavior change theory. Social support, cognitive, and self-regulatory theories were the only models used in this category [35,45,52,72-74]. The analysis showed 50% (3/6) of video-game platforms used behavior change theories, where the cognitive and self-regulatory theories are the only used [75,76]. Only 11% (2/18) of telemonitoring studies used biomedical and cognitive theories [77,78]. Literature showed that 89% (71/80) of studies with behavior theories had a significant impact on health outcomes. Only 11% (9/80) of telemonitoring studies explicitly referenced the use of behavior theories and showed no impact of technology on health outcomes. The result failed to show any relationship between using behavior theory and the impact of technology on health outcomes ($\chi^2=0.008, P=.977$).

The analysis also found no significant correlative relationship between behavior theory and patient engagement in health care ($\chi^2=0.3055, P=.580479$). However, there was a significant relationship between patient engagement and Internet-based interventions using behavior theories ($\chi^2=7.3144, P=.00684$) (see Table 5).
Table 5. Patient engagement and theories of health behavior (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Patient engagement, n (%)</th>
<th>Behavior theory</th>
<th>Video game</th>
<th>Total</th>
<th>Total no</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internet</td>
<td>Mobile</td>
<td>Social media</td>
<td>Tele-monitoring</td>
</tr>
<tr>
<td>Yes</td>
<td>2 (1)</td>
<td>1 (6)</td>
<td>2 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>No</td>
<td>1 (1)</td>
<td>1 (6)</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>1 (6)</td>
<td>2 (1)</td>
<td>2 (1)</td>
<td>2 (1)</td>
</tr>
</tbody>
</table>

Methods to Measure Health Outcomes

Most studies used indirect ways (such as self-reports) to measure health outcomes (65.9%, 112/170). The literature showed that 57.6% (98/170) of studies showed a positive impact of IT platforms when the health outcomes were assessed using indirect ways. For example, self-reporting was used to assess whether a text message could increase smoking cessation [68], reduce methamphetamine use among human immunodeficiency virus patients [79], and to assess medication adherence among patients with congestive heart failure [48]. The analysis showed no significant association between ways to measure health outcomes and technology impact ($\chi^2=0.5793$, $P=.446603$) (see Table 6).

Table 6. Impact of IT platforms and methods to measure health outcomes (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Impact of information technology platforms, n (%)</th>
<th>Methods to measure health outcomes</th>
<th>Internet</th>
<th>Mobile</th>
<th>Social media</th>
<th>Tele-monitoring</th>
<th>Video game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total yes</td>
<td>Direct</td>
<td>25 (29)</td>
<td>3 (3)</td>
<td>18 (41)</td>
<td>2 (5)</td>
<td>6 (33)</td>
</tr>
<tr>
<td>Total no</td>
<td>Direct</td>
<td>3 (50)</td>
<td>53 (31)</td>
<td>5 (3)</td>
<td>58 (34)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Indirect</td>
<td>50 (58)</td>
<td>8 (9)</td>
<td>23 (52)</td>
<td>1 (2)</td>
<td>10 (56)</td>
</tr>
<tr>
<td></td>
<td>Indirect</td>
<td>3 (50)</td>
<td>98 (58)</td>
<td>14 (8)</td>
<td>112 (66)</td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>Indirect</td>
<td>75 (87)</td>
<td>11 (13)</td>
<td>41 (93)</td>
<td>3 (7)</td>
<td>16 (89)</td>
</tr>
<tr>
<td></td>
<td>Grand Total</td>
<td>75 (87)</td>
<td>11 (13)</td>
<td>41 (93)</td>
<td>3 (7)</td>
<td>16 (89)</td>
</tr>
</tbody>
</table>

Usability Assessment

Only 33.5% (57/170) of studies assessed the usability of IT platforms. Of those, the majority were considered by authors to be usable (89%, 51/57). Specifically, 75% (28/37) of Internet-based IT intervention studies showed positive health outcomes with usable interventions [41,80-106]. In one study that gauged usability, Steele et al performed a 3-month randomized controlled trial among 192 participants and found an Internet-based physical activity behavior change program to be usable, feasible, and acceptable among inactive participants [41]. Mobile-based interventions also showed 75% (6/8) of usable interventions had a positive impact on health outcomes [42,69,93,107-109]. In one study, SMS was found to be useful in helping patients to remember to take their medications and be engaged in treatment planning [107]. SMS-based intervention was also found to be useful in promoting communication with health care providers by delivering, receiving health information, generating questions, and seeking information related to health conditions [42]. Moreover, 87% (7/8) of studies reported that the usability of social media-based interventions was positively correlated with good impact on health outcomes [34,35,46,52,53,110,111]. One particular social networking-related study found that online health-related social networking was useful and acceptable in chronic disease management [52]. In addition, one study reported the usability assessment in the video-game category and found that it was usable and had a positive impact among patients with hyperfunctional voice disorders [36]. Overall, the analysis also found that telemonitoring also showed similar results (100%, 3/3). One telemonitoring-based study found that
telecommunication-based reminder tools are useful for improving medication adherence [112].

Although our results failed to report any relationship between usability of IT platforms and the impact on health outcomes (P= .1065), they showed significant association between usability and patient engagement in health care (P= .0216) (Fisher’s exact test) (see Tables 7 and 8).

Table 7. Impact of IT platforms and usability (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Usability</th>
<th>Impact of information technology platforms, n (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internet</td>
<td>Mobile</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Usable</td>
<td>28 (33)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Not usable</td>
<td>1 (1)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Total of assessed usability</td>
<td>29 (34)</td>
<td>8 (9)</td>
</tr>
<tr>
<td>Not assessed usability</td>
<td>46 (53)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Grand total</td>
<td>75 (87)</td>
<td>11 (13)</td>
</tr>
</tbody>
</table>

Table 8. Patient engagement and usability (Yes=positive impact, No=no impact).

<table>
<thead>
<tr>
<th>Usability</th>
<th>Impact of information technology platforms, n (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internet</td>
<td>Mobile</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Usability assessed (usable)</td>
<td>26 (30)</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Usability assessed (not usable)</td>
<td>1 (1)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Total usability assessed</td>
<td>27 (31)</td>
<td>10 (12)</td>
</tr>
<tr>
<td>Not assessed</td>
<td>41 (48)</td>
<td>8 (9)</td>
</tr>
<tr>
<td>Grand total</td>
<td>68 (79)</td>
<td>18 (21)</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

Impact of IT Platforms on Health Outcomes

Overall, this review indicated that IT platform-based health interventions had a great impact on patients’ health outcomes in the United States and in other nations. IT-based health interventions have been viewed as driving positive health behavior change through patient engagement with most technology platforms. IT-based health interventions also provide necessary information and advice and counseling related to certain diseases and conditions, such as mental disorders [113-120], asthma [57,76,96,121-123], obesity [30,32,124-132], smoking [40,68,69,88,133-136], diabetes [11,137-151], sleep disorder [152], hypertension [127,153,154], cancer [34,58,60,74,75,82,92,97,98,155-157], thereby encouraging healthy living [30,31,72,124,158,159]. Moreover, these interventions enable patients to be engaged in self-monitoring, thereby directing patients toward healthy eating, enhancing attendance rate [136,160,166], improving medication adherence [162,167-173], increasing knowledge about disease and treatment [33,42,47,75,85,90,94,96,115,119,168,174,175], and enhancing exercise use [32,56,80-82,122,125,126,128-132,153,156,176-183]. Online coaching by specialists enables patients to recover quickly, ensuring that the pain they experience is reduced [89,184], and doctor-patient communications are made readily available [73,157,185-191].

Apart from Internet-based technologies, mobile phone technologies have been used extensively to engage patients and ensure there is patient health behavior change. Mobile phone technologies engage patients by using SMS to contact them and provide necessary health information. This technology can be very effective and efficient, since it is less expensive and therefore more people can afford it. Studies have shown that patients can receive health-related information, receive reminders of their health care attendance, as well as be encouraged to adhere to their treatment [35,77-67,107,192].

Social media outlets, such as Twitter and Facebook, can ensure patients get and exchange necessary health information [34,46]. Video game and telemonitoring technologies served a similar purpose; these technologies tried to engage patients in order to provide necessary health information and provided a platform for helping patients adhere to treatment and helped patients...
actively become involved in the treatment process. These technologies are of great importance to patients as well as helpful to health care providers, therefore ensuring effectiveness and efficiency.

Although several studies demonstrated the positive impact of IT platform usage, others showed no impact [33,35,41,43,45-48,81,83,91,133,141,178,193-198]. This could be due to the timing of the follow-up assessments ranging from one extremely short follow-up timing (1 week) to a relatively long-term follow-up timing (48 months). The lack of consistency in follow-up timing made it unclear as to how long these effects on patient health last. Moreover, the technology adoption rate may decline after a certain time period, thus diminishing its effectiveness after significant results at the beginning of the study. This occurred in a study by Williamson et al who found that after 2 years of an IT-based intervention, the decrease in body weight did not differ between the intervention and control group [28]. Similarly, another research study found a slow decline in HbA1c at 3 months follow-up (1.22%) versus (1.09%) 6 months follow-up [199,137]. Therefore, designing and evaluating IT platforms may become a significant challenge because researchers are dealing with a large volume of interventions that have different impacts on patient health behavior. Thus, several issues need to be addressed if such interventions are to be evaluated or assessed, such as length of intervention, type of technology, usability of the technology, application of behavior theory, and how health outcomes are measured.

**Patient Engagement in Health Care Using IT Platforms**

Our review showed that IT platforms are playing a significant role in patient engagement. This review implies that higher patient participation in condition self-management was correlated with greater improvement in health outcomes. Many studies have shown that patients who actively participated in health care experience better health outcomes compared to less involved patients. One specific study showed a significant association between patient engagement using the Internet and weight loss at 6 months ($P<.001$) [77]. Another study reported that a text messaging-based intervention could enhance patient engagement [33]. Social networks can also be particularly helpful to individuals with lower patient activation [52]. Despite the evidence regarding the importance of patient engagement, it is challenging to draw solid conclusions. Many of the studies conducted qualitative surveys to measure patient engagement or relied solely on the number of times patients logged in or uploaded data to determine their engagement. However, system log-ins and upload and download data are not engagement. Patient engagement is basically about interaction and participation in managing one’s health to achieve desired goals. Therefore, further research is needed to determine the best ways to measure patient engagement.

**Association Between Usability of IT Platforms and Their Impact**

Our review found limited levels of evidence supporting the correlation between usability and impact of technology on health outcomes ($P=.1065$). Several factors may hinder the positive impact of technology on health outcomes other than usability issues. Patients’ willingness to participate in managing their health care could be one of the main reasons. The review found a significant relationship with patient engagement and impact of technology. It found also a significant positive correlation between patient engagement and usability of IT platforms. Even though the aim of this study was not to discover determinants of patient engagement, several issues were identified including unequal access to technology, technical issues, poor interface design, suboptimal message content, privacy and confidentiality issues [46,108,110,121,165,193,200], and patients’ self-perceived health illiteracy. The latter issue was seen in social media, where patients think such a discussion should be restricted to health care professionals [46]. Also, the majority of technologies rely on patient-provider engagement from both sides to exchange information and manage health conditions, such as in two-way SMS, thus increasing burden on providers as well as patients. Moreover, in some countries like Sweden, information dissemination can be restricted by legal and ethical regulations for online patient-provider communication [201]. Therefore, more research on the usability and acceptability of these technologies and discovering the different factors that impact patient engagement and their meaningful use will be required in the future.

**Association Between Technology Impact and Intervention Grounded in Behavior Change Theories**

This review found that only a limited number of specific behavioral theories and models were referenced among multiple articles inferring a theoretical design. This could imply that several IT interventions are designed in an ad hoc way, without using any theoretical frameworks. This finding supports the results of a previous study showing the majority of mobile-based interventions used for improving medication adherence and disease management were developed without a theoretical basis [202]. The review failed to detect any relationship between (1) behavioral theories and impact of technology or (2) theories and patient engagement. This could imply that existing theories/model were not developed to be used with these technologies. We found a significant association between patient engagement in Internet-based interventions and use of behavior theories in these interventions ($\chi^2=7.3144$, $P=.00684$). This could imply that existing theories or models may have limited applicability. However, it was difficult to draw a clear conclusion whether or not using theory influenced intervention effectiveness. Possible reasons for the lack of theory may include the investigator not citing the theory, researchers’ lack of knowledge of the theories, struggling to define appropriate theories, poorly operationalized theories, an absence of good evaluation methods and usability testing, and theories containing overlapping constructs and inconsistent use of terminology. For example, the construct of self-efficacy can be found in Social Cognitive Theory, Protection Motivation Theory, the Theory of Planned Behavior, the Health Belief Model, and Self-Regulation Theory. In addition, the simplicity of the interventions could be another reason for not including behavior theories. For example, reminding patients to take their treatment through text message appears simple and consistent with the “cue to action” constituent of many health behavior theories or models, but these theories were not always described. Our
findings of the lack of association between use of theory and outcomes was based on the theory description within each published article and should be interpreted cautiously.

**Association Between Methods to Measure Health Outcomes and the Impact of These Technologies**

Overall, slightly more than half of the reviewed articles had a positive impact when assessed with patient questionnaires, patient self-reports, pill counts, rates of prescription refills, assessment of patients’ clinical response, and electronic medication monitors. Even though the way to measure health outcomes is an important factor in determining the impact of technology, the review failed to detect any relationship between methods used to measure health outcomes and the impact of technology. Therefore, further study is needed to replicate our results, because for each approach, there are different assumptions related to what data to collect, how to collect that data, and how to make decisions about success. Indirect methods may overestimate patient adherence. For instance, metformin treatment adherence can be monitored either by recording the number of times the medication bottle was opened, or alternately, adherence could be gauged by metformin plasma levels. Both health behaviors are part of the same behavioral class to control blood sugar levels. However, measuring metformin in blood is more effective at measuring adherence than recording the time when the bottle is opened because patients may open and close the bottle without taking any medication.

**Limitations**

Our review included some limitations. First, due to the heterogeneity of the research studies and the fact that some data were not available for certain types of interventions and their characteristics, some statistical tests could not be performed, hindering optimal quantitative assessment. Second, we excluded studies not written in English; this criterion might have omitted certain relevant research. Third, the majority of studies were performed in the United States, which limits generalizability of findings. Finally, because of possible publication bias toward positive findings, our review may overestimate the actual impact of these technologies.

**Implications**

The results from this review reveal several practical applications worthy of future study (summarized in Table 9).
Table 9. Implications of study.

<table>
<thead>
<tr>
<th>Suggestion</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information technology platforms</td>
<td>It would be valuable to further evaluate IT platform-based interventions to form a more coherent picture of their effectiveness in encouraging patient engagement for the purpose of enhancing lasting health behavior change. A study with a long time frame may be useful to draw a clear conclusion on the effectiveness of these technologies and to determine the best ways to guarantee positive long-term effects in patients. Also, due to low availability of studies meeting our criteria, we could not provide or conclude relationships between factors. Therefore, we recommend doing another review when there are more studies available in future. In future, we can increase the quality of the review by limiting sample size and study time frame. IT platform interventions reviewed in this study are mutually inclusive; they use different labels and contexts to describe the same concepts and lack of formal definitions. Therefore, a common framework for analyzing these concepts is needed. A framework with an ontological approach may serve this purpose.</td>
</tr>
<tr>
<td>Patient engagements</td>
<td>The outreach and engagement period prior to the intervention enrollment are critical to the success of any intervention. Therefore, studies should consider that when implementing the interventions A study assessing determinant of patient engagement is highly recommended.</td>
</tr>
<tr>
<td>Usability</td>
<td>Assessment of user satisfaction toward IT platforms and their usability of these platforms are needed, and could be done through qualitative evaluations of user opinions of the respective IT platform(s).</td>
</tr>
<tr>
<td>Theories of health behavior</td>
<td>The literature also needs to focus more on referencing, selecting, and implementing behavioral theory to achieve the best possible impact. Reporting accurate information about interventions is essential to assessing the effectiveness of these interventions and facilitating their successful implementation. Also, new theories are needed to better understand how patients can participate and facilitate health behavior change, theories building on past conceptual and focus only on one aspect, a triangulation model would provide internally logical and comprehensible perception to achieve these goals.</td>
</tr>
<tr>
<td>Methods measure health outcomes</td>
<td>It would be valuable to further examine how different types of measurement could affect patient outcomes reported in the study. A comparison between direct and indirect methods could be helpful to draw a clear conclusion.</td>
</tr>
</tbody>
</table>

Conclusions

Based on our review, there is moderately strong evidence that IT platforms can engage patients in health care and improve health outcomes. The usefulness and acceptability of IT platforms can have great power in engagement and outcomes. Studies grounded in behavior theory appeared to show a positive impact on patient health behavior. To exploit the full potential of IT platforms in health care, new theories may be needed to better understand how patients can participate and facilitate health behavior change. Selecting appropriate ways to measure health behavior change and developing a common framework to analyze and understand the different components of IT platforms and their safety, effectiveness, efficiency, and acceptability will also be of great importance.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Quality Assessment of Included Articles.

[PDF File (Adobe PDF File), 137KB - medinform_v4i1e1_app1.pdf]
Multimedia Appendix 2

List of the included articles and their characteristics.

[PDF File (Adobe PDF File), 233KB - medinform_v4i1e1_app2.pdf]

Multimedia Appendix 3

A glossary of terms.

[PDF File (Adobe PDF File), 107KB - medinform_v4i1e1_app3.pdf]

References


10. Martyn H, Gallant LM. Over 50 and wired: Web-based stakeholder communication. FM 2012 Jun 08;17(6) 17(6) [FREE Full text] [doi: 10.5210/fm.v17i6.3449]


Abbreviations

IT: information technology
SMS: short message service
SOC: stage of change
TPB: theory of planned behavior
TTM: transtheoretical model

Edited by G Eysenbach; submitted 09.04.15; peer-reviewed by R Schnall, L Morrison, N Kaufman; comments to author 29.07.15; revised version received 07.09.15; accepted 09.10.15; published 21.01.16

Please cite as:
Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF
URL: http://medinform.jmir.org/2016/1/e1/
doi: 10.2196/medinform.4514
PMID: 26795082

©Suhila Sawesi, Mohamed Rashrash, Kanitha Phalakornkule, Janet S Carpenter, Josette F Jones. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 21.01.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.
Early Indication of Decompensated Heart Failure in Patients on Home-Telemonitoring: A Comparison of Prediction Algorithms Based on Daily Weight and Noninvasive Transthoracic Bio-impedance

Ilapha Cuba Gyllensten1,2, MSc; Alberto G Bonomi1, PhD; Kevin M Goode3, PhD; Harald Reiter1, MSc; Joerg Habetha1, PhD; Oliver Amft2,4, PhD; John GF Cleland5, MD, PhD

1Personal Health Solutions, Philips Research, Eindhoven, Netherlands
2Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands
3Department of Health Professional Studies, Faculty of Health & Social Care, University of Hull, Kingston-Upon-Hull, United Kingdom
4ACTLab, University of Passau, Passau, Germany
5National Heart & Lung Institute, Imperial College, London, United Kingdom

Corresponding Author:
Ilapha Cuba Gyllensten, MSc
Personal Health Solutions
Philips Research
5.007
High Tech Campus 34
Eindhoven, 5656AE
Netherlands
Phone: 31 631926930
Fax: 31 40274276
Email: illapha@gmail.com

Abstract

Background: Heart Failure (HF) is a common reason for hospitalization. Admissions might be prevented by early detection of and intervention for decompensation. Conventionally, changes in weight, a possible measure of fluid accumulation, have been used to detect deterioration. Transthoracic impedance may be a more sensitive and accurate measure of fluid accumulation.

Objective: In this study, we review previously proposed predictive algorithms using body weight and noninvasive transthoracic bio-impedance (NITTI) to predict HF decompensations.

Methods: We monitored 91 patients with chronic HF for an average of 10 months using a weight scale and a wearable bio-impedance vest. Three algorithms were tested using either simple rule-of-thumb differences (RoT), moving averages (MACD), or cumulative sums (CUSUM).

Results: Algorithms using NITTI in the 2 weeks preceding decompensation predicted events ($P<.001$); however, using weight alone did not. Cross-validation showed that NITTI improved sensitivity of all algorithms tested and that trend algorithms provided the best performance for either measurement (Weight-MACD: 33%, NITTI-CUSUM: 60%) in contrast to the simpler rules-of-thumb (Weight-RoT: 20%, NITTI-RoT: 33%) as proposed in HF guidelines.

Conclusions: NITTI measurements decrease before decompensations, and combined with trend algorithms, improve the detection of HF decompensation over current guideline rules; however, many alerts are not associated with clinically overt decompensation.

(JMIR Med Inform 2016;4(1):e3) doi:10.2196/medinform.4842

KEYWORDS
Heart failure; telemonitoring; deterioration detection; alert algorithms; ambulatory monitoring; impedance
**Introduction**

Chronic heart failure (HF) is common [1] and a substantial drain on scarce healthcare resources [2]. Much of the costs of HF are due to the high rate of unplanned admissions for worsening HF. For patients who survive an admission for worsening HF, rehospitalization rates are high and >20% will die within one year [3,4]. Furthermore, the high prevalence and costs associated with HF are projected to rise as the population ages [5]. Telemonitoring could reduce costs and improve outcomes [6] by substituting infrequent assessments at a clinical facility by a health professional with frequent remote monitoring done by patients themselves. This could facilitate more timely and tailored interventions. The efficacy of telemonitoring would be greatly improved if decompensation events could be detected before the onset of severe symptoms [7,8].

Worsening heart failure may lead to weight gain as a consequence of fluid retention and edema and, if untreated, can lead to hospitalization and ultimately death. The Heart Failure Association of America (HFSA) and the European Society of Cardiology (ESC) guidelines both recommend daily weight monitoring. The ESC recommends that patients experiencing a weight increase of 2 kg or more in 3 days should alert healthcare professionals and increase their diuretic dose [9]. The HFSA recommends the restriction of sodium and water after an increase of more than 2 lbs (0.9 kg) in 1 day, or more than 4 lbs (1.8 kg) over a week, followed by an alert to healthcare professionals if the increase continues [10].

Worsening hemodynamics with increased vascular resistance, afterload mismatch, congestion, and diastolic dysfunction are thought to precede fluid accumulation [11]. Increased end-diastolic pulmonary arterial pressure (PAP), a direct measure of hemodynamic overload, and decreased intrathoracic impedance (ITI), an indirect measure of pulmonary congestion, have both been observed in the days and weeks prior to decompensation [12-14]. Thoracic impedance can also be measured noninvasively (NITTI) [15], which correlates with ITI [16], making measurement possible in a far broader range of patients. NITTI measures a much larger field; however, the variability in measurements may depend on the patients’ willingness and ability to position electrodes accurately. Recently, several new wearable devices have been proposed for this purpose, such as specialized vests [17,18] or adhesive patches [19,20].

An increased risk of decompensation has been shown for both weight gain [21] and decline in ITI [22]; however, recent studies have shown that absolute changes in weight over short time periods are not sensitive in detecting impending decompensation [23-25], and that ITI may have high sensitivity but a high rate of false alarms per patient-year [26]. However, to the authors’ knowledge, recently proposed prediction algorithms comparing body weight and impedance head-to-head have not been investigated using noninvasive technology.

The aim of this investigation was to evaluate and compare the predictive value of previously published algorithms using measurements of daily body weight, and noninvasive measures of NITTI from a smart-textile vest, to detect decompensation prior to the onset of severe symptoms leading to hospitalization.

**Methods**

**Patient Population**

The data for this analysis were collected as part of the MyHeart heart failure management observational study [27]. The MyHeart study was unique in its collection of several different vital signs and innovative markers using noninvasive sensors and a home-telemonitoring system. Six HF clinics in Germany and Spain participated in the collection of the clinical data. Patients were included in the study if they had chronic HF with an elevated N-terminal of the prohormone brain natriuretic peptide (NT-proBNP ≥ 500 pg/ml), were taking at least 40 mg/day of furosemide or an equivalent, and were in the New York Heart Association (NYHA) functional class II, III, or IV. They were excluded if they had the following: severe chronic obstructive pulmonary disease (COPD GOLD Class > 2), primary pulmonary hypertension, renal insufficiency requiring dialysis, a psychiatric or neurological disorder of moderate to severe degree (eg, dementia, schizophrenia, substance disorder, psychotic depression), prior acute myocardial infarction or coronary artery bypass grafting (CABG) in the previous 3 months. Ethical approval was provided by the Medical Ethics Committees in the 2 respective countries.

Of 148 patients recruited from October 2008 to July 2010, 108 had the system installed and data recorded; 3 did not fit the criteria, 3 were unavailable at installation, 1 died before installation, and 33 withdrew before system installation. Of the remaining 108 users, 17 used the system on less than 30 occasions, leaving 91 patients as the focus of this exploratory analysis. Their mean (SD) age was 63 (12) years and 64 were men. Mean weight was 84 (19) kg, mean BMI was 29 (6) kg/m², and mean left ventricular ejection fraction (LVEF) was 31 (12)%.

Most patients had mild (NYHA class II: 60%) or moderate (NYHA class III: 36%) symptoms. Etiology was ischemic in 47%, idiopathic dilated cardiomyopathy in 31%, valvular disease in 5%, and other in 9%. Comorbidities included hypertension (68%), diabetes (37%), atrial fibrillation (36%), renal dysfunction (28%) and COPD (13%). Treatment included angiotensin converting enzyme (ACE) or angiotensin receptor blockers (ARB) (87%), beta-blockers (88%), MRA (53%), diuretics (84%), digoxin (21%), and implantable cardioverter-defibrillator/cardiac resynchronization therapy (ICD/CRT) (23%/14%). The average monitoring time was 10 months, during which 19 patients were hospitalized one or more times due to decompensated HF, with a total of 24 decompensated HF hospitalizations. The adverse events were adjudicated by an advisory committee.

**Daily Measurements of Body Weight and NITTI**

Patients were instructed on how to perform measurements of body weight and NITTI. Measurements were carried out in the morning before eating breakfast. Body weight was collected using a weight scale (Philips Medical Systems, Andover, Massachusetts, USA), which automatically logged the measurements (accuracy ± 0.1 kg). TTI was measured using a...
wearable bio-impedance vest [28], shown in Figure 1. The vest
measures TTI at several electrical frequencies (10 kHz-1MHz).
These recordings give a characterization of the electrical
properties of the tissue, as described by the Cole-Cole model
[29]. At low measurement frequencies, biological tissue
impedance is mainly determined by the extracellular fluid
content and characteristics. At higher frequencies, electrical
properties are determined by both the intracellular and
extracellular fluid content. Multi-frequency measurements of
thoracic bio-impedance therefore allow isolation of the Cole
parameters that indirectly reflect either the intracellular or
extracellular fluid content. We used the external resistance
derived from the Cole-Cole model, since this indirectly reflects
extracellular water, which is the component associated with
decompensation. In another study, we have shown that this
metric tracks changes in symptoms and fluid loss during
treatment for decompensated HF [17].

Figure 1. The bioimpedance vest shown by a model subject correctly applying it across the chest. Textile electrodes on each side of the flexible
measurement panel inject currents at different frequencies and register the resulting voltage to calculate the impedance parameter relating to extracellular
fluid volume.

Alarm and Event Definition
The weight and NITTI data were applied to published algorithms
detailed description in Multimedia Appendix 1), to predict the
onset of decompensation prior to subsequent hospitalization
due to worsening heart failure. The output of these algorithms,
the output index, could be as simple as the difference between
the current measurement and the measurement made 2 days
previously, or a more complex calculation (eg, one based on
cumulative sums). An alert is triggered when the output index
exceeds a specific threshold.

The predictive power of the algorithms was assessed by
exploring their ability to alarm within a prespecified period
before a hospitalization due to worsening heart failure. Changes
in NITTI are thought to precede changes in weight prior to
hospitalization [12,21]. Depending on the measure used,
previous studies have considered alerting periods from 2 weeks
[23] up to one month [30] before hospitalization. In this study,
a 2-week period was chosen as an adequate period before a
hospitalization, during which alarms should be raised, giving
time for the patient or clinician to act. Alerts occurring outside
of this period were counted as false alarms. Short periods of a
few days at the start of monitoring, end of monitoring, and
directly following a hospitalization did not fit into any 2-week
division and were subsequently removed from the analysis.

Performance Assessment of Algorithms
Three types of alert algorithms are compared in this study:
rule-of-thumb (RoT) [21,23,26], moving average convergence
divergence (MACD) [23], and cumulative sum control chart
(CUSUM) [31]. The qualitative differences between these are
shown in Figure 2. Rule of thumb (RoT) methods provide a
noisy measure for which chance readings have a large effect,
sometimes with no underlying trend; however, they also provide
a fast response to changes. Moving averages (MACD) react
more slowly but follow underlying trends better, in both
directions. Cumulative sums (CUSUM) provide uni-directional
detection and lead to longer sustained alerts. For a detailed
description of the definitions of each algorithm see Appendix
1. The predictive performance of the algorithms was compared
using receiver operator curve (ROC) analysis. The sensitivity
and specificity of each algorithm was calculated by dividing
the measurement data into periods of 2 weeks, in such a way
that a period containing a decompensated hospitalization would
end when the hospitalization event occurred. This led to the
following definitions:

1. True positive: An alarm during the 2 weeks preceding a
hospitalization;
2. False positive: An alarm during any other 2-week period;
3. True negative: A 2-week period without any alarms;
4. False negative: A 2-week period ending in a hospitalization
without any alarms.
Algorithm Selection and Optimization

Each of the algorithms considered in this study (RoT, MACD, CUSUM) have modifiable parameters that will alter their behavior and ultimately their predictive performance. We tested the performance of each algorithm for a range of possible parameter values. For the RoT algorithms, the number of days \( d \) between the measurements used to calculate the difference was varied from 1 day to 21 days. In the MACD algorithm, the long-term average parameter \( N_l \) was varied between 10 and 50 days in increments of 5 days, and the short-term moving average parameter \( N_s \) was varied between 1 and 10 days. In the CUSUM algorithm, the parameter determining the length of the running mean and standard deviation \( d \) was varied between 10 and 30 days in increments of 5 days, and the parameter determining the depreciation of the accumulated sum \( c \) was evaluated between 0.5 and 1.5, in increments of 0.2.

Segmentation of the data into 2-week periods results in substantially more periods without an HF-related hospitalization compared to those with one. To avoid producing algorithms that raise a large number of false positive alarms, previous studies have focused only on alarms with high specificity \([20,30]\). In this investigation, the best parameters were chosen to be those that maximized the area under the curve for thresholds with a specificity >95%. The output index for each algorithm was then normalized to allow the correct estimation of the ROC curves during the cross-validation procedure described below.

Parameter optimization can lead to models that overfit the data, which then would not generalize well to other data sets. To minimize these effects, we implemented a stratified leave-patient-out cross-validation (CV) method for the parameters in the RoT, MACD, and CUSUM algorithms. This procedure randomly splits the data into 8 groups, while maintaining the number of patients and decompensation events in each group. The parameters were then optimized for the data with one group left out. The data from the left-out group were then used to evaluate the performance of the optimized parameters. This was repeated until all groups had been left out once. The left-out groups were then recombined to provide an unbiased ROC curve. The optimal threshold for the output index was chosen to be the Youden point with specificity larger than 90%.

Statistics

Comparisons between the recorded measurements and the output index for the different algorithms in the 2 weeks preceding hospitalization and all other periods were tested with a mixed-effect model using patient specific intercepts as random effects. An arbitrary significance of 0.05 was assumed throughout. Missing data due to adherence issues were removed from the analysis by excluding periods in which less than 3 \([32]\) measurements per week were found. In the case of algorithms that needed previous data points to estimate trends, a linear imputation between adjacent data points was carried out. It should be noted that when the algorithms processed the data, imputations were only made on data that would have been available for a system running in real time; no imputations using future values were done. NITTI measurements were log-transformed to adjust for skewness. All listed algorithms were developed and evaluated using the software suite MATLAB 7.13.0.564.
**Results**

**Data Characteristics**

Among the 91 patients for whom data were included in the analysis, 24 heart failure-related hospitalizations occurred in 19 patients. Of the 24 hospitalizations, 9 had less than 3 weekly weight recordings and 12 had less than 3 weekly impedance recordings preceding the hospitalization, and were excluded from the analysis. The minimum window for the CUSUM algorithm excluded an additional 2 for its analysis.

**Prediction Performance**

The predictive performance of guideline-based rules and published algorithms using weight are presented in Table 1. With the exception of those rules with very low specificity (ie, <60%), all rules based on short-term increases had low sensitivity when applied to the data (typically <25%). Rules based on longer-term increases showed higher sensitivity; however, only one had a specificity >90%. The MACD algorithm with the parameter proposed by Zhang et al. [23] outperformed the other weight algorithms.

The cross-validation analyses of the developed models based on published algorithms are presented in Figure 3. The RoT-based algorithms using weight have poor sensitivity at a specificity between 90-100%, with performance close to random chance. This poor sensitivity was also observed when evaluating previous published guidelines using windows between 2 and 3 days (Table 1). As expected, this sensitivity increased when longer windows and/or lower thresholds were used, but at the cost of a lower specificity.

The MACD algorithm improved performance for both weight and impedance. The CUSUM algorithm improved performance for NITTI. The performance of trend algorithms was superior to previously published algorithms (Table 1).

**Table 1.** Performance of different weight algorithms in anticipating an upcoming decompensation.

<table>
<thead>
<tr>
<th>Source</th>
<th>Weight algorithm</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV&lt;sup&gt;a&lt;/sup&gt; (%)</th>
<th>NPV&lt;sup&gt;b&lt;/sup&gt; (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guideline issuing bodies</td>
<td>&gt;2 lbs&lt;sup&gt;c&lt;/sup&gt; in 1 day [10]</td>
<td>67</td>
<td>56</td>
<td>1.4</td>
<td>99.5</td>
</tr>
<tr>
<td></td>
<td>&gt;2 kg in 3 days [9]</td>
<td>13</td>
<td>87</td>
<td>0.9</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;4 lbs&lt;sup&gt;c&lt;/sup&gt; in 1 week [10]</td>
<td>27</td>
<td>87</td>
<td>1.8</td>
<td>99.2</td>
</tr>
<tr>
<td>Existing literature</td>
<td>Random chance</td>
<td>50</td>
<td>50</td>
<td>0.9</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;2 lbs in 1 day or &gt;3 lbs in 3 days [26]</td>
<td>73</td>
<td>50</td>
<td>1.3</td>
<td>99.5</td>
</tr>
<tr>
<td></td>
<td>&gt;2 lbs in 1 day or &gt;5 lbs in 3 days [26]</td>
<td>67</td>
<td>56</td>
<td>1.4</td>
<td>99.4</td>
</tr>
<tr>
<td></td>
<td>&gt;3 lbs in 1 day or &gt;5 lbs in 3 days [26]</td>
<td>13</td>
<td>82</td>
<td>0.7</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;3 lbs in 1 day or &gt;7 lbs in 3 days [26]</td>
<td>7</td>
<td>83</td>
<td>0.4</td>
<td>99.0</td>
</tr>
<tr>
<td></td>
<td>&gt;4 lbs in 1 day or &gt;7 lbs in 3 days [26]</td>
<td>7</td>
<td>93</td>
<td>0.9</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;4 lbs in 1 day or &gt;9 lbs in 3 days [26]</td>
<td>7</td>
<td>93</td>
<td>0.9</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;5 lbs in 1 day or &gt;9 lbs in 3 days [26]</td>
<td>0</td>
<td>100</td>
<td>—</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td>&gt;2 lbs in 1 week [21]</td>
<td>80</td>
<td>45</td>
<td>1.3</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td>&gt;5 lbs in 1 week [21]</td>
<td>20</td>
<td>94</td>
<td>2.7</td>
<td>99.2</td>
</tr>
<tr>
<td></td>
<td>&gt;4 lbs in a 5 to 80 days MACD&lt;sup&gt;d&lt;/sup&gt; [23]</td>
<td>20</td>
<td>97</td>
<td>6.3</td>
<td>99.3</td>
</tr>
</tbody>
</table>

<sup>a</sup>PPV: positive predictive value  
<sup>b</sup>NPV: negative predictive value  
<sup>c</sup>To convert to kilograms multiply by 0.45  
<sup>d</sup>MACD: moving average convergence divergence
**Figure 3.** ROC curves from the cross-validated evaluation for the three considered algorithms in the specificity range from 0.9 to 1. A shows the rule of thumb algorithm, B the MACD algorithm, and C the CUSUM algorithm. Performance using NITTI measures is shown with the dashed green line, weight is shown with the blue line, and random chance is portrayed by the red dotted line.

**Optimal Parameters**

The output of the 2 best performing algorithms for weight and impedance with optimal parameters (maximum Youden index with specificity >90%) is shown in Figure 4. Clear trends in both weight and impedance can be seen for Patient 1 and both algorithms managed to alert before the decompensation; a full week in advance for impedance and a day in advance for weight. Patient 2, on the other hand, had no or weakly visible trends, which were not enough to trigger an alert. The patient did exhibit large daily weight fluctuations, which could have indicated instability; however, this was not picked up by the algorithms. The optimal parameters for all 3 algorithms for weight and impedance are shown in Table 2, together with the cross-validated performance measures. Both trend algorithms using NITTI outperformed the weight algorithms.

**Table 2.** Cross-validated performance measures of the algorithms at the maximum Youden index within a specificity of 90-100%.

<table>
<thead>
<tr>
<th>Optimal algorithms</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>PPV b</th>
<th>NPV c</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weight</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RoT d: &gt;2.7 kg in 17 days</td>
<td>20</td>
<td>90</td>
<td>1.95</td>
<td>99.2</td>
</tr>
<tr>
<td>MACD e: &gt;0.62 kg (N_s=9, N_l= 20 days)</td>
<td>33</td>
<td>91</td>
<td>3.2</td>
<td>99.3</td>
</tr>
<tr>
<td>CUSUM f: &gt;8.7 with 10-day average, c=0.75</td>
<td>13</td>
<td>91</td>
<td>1.4</td>
<td>99.1</td>
</tr>
<tr>
<td><strong>NITTI g</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RoT: &lt;-0.27 (log ohm) in 21 days</td>
<td>33</td>
<td>92</td>
<td>4.2</td>
<td>99.2</td>
</tr>
<tr>
<td>MACD: &lt;-0.059 (log ohm) (N_s=9, N_l= 35 days)</td>
<td>50</td>
<td>92</td>
<td>5.9</td>
<td>99.5</td>
</tr>
<tr>
<td>CUSUM: &lt;-7.8 with 20-day average, c=0.75</td>
<td>60</td>
<td>96</td>
<td>10.9</td>
<td>99.6</td>
</tr>
</tbody>
</table>

*The optimal parameters and thresholds were estimated from the full data (for stability and variance of cross-validated parameters and thresholds, see Table 3).*  
*PPV: positive predictive value*  
*NPV: negative predictive value*  
*RoT: rule of thumb*  
*MACD: moving average convergence divergence*  
*CUSUM: cumulative sums*  
*NITTI: noninvasive transthoracic bio-impedance*
Figure 4. Three weeks of telemonitoring data from two patients with high compliance before an upcoming decompensation. Circles correspond to NITTI measurements and the NITTI-CUSUM algorithm and crosses correspond to weight measurements and the weight-MACD algorithm. Optimal thresholds are shown as dash-dotted lines in green for NITTI and dotted blue lines for weight.

Algorithmic Stability
The use of a cross-validation procedure to minimize biased performance measures generated several plausible parameters for the tested algorithms; these are presented in Table 3. In general, RoT had lower variance in estimated parameters than MACD, which in turn had lower variance than CUSUM, coinciding with the increasing complexity of the algorithms. Parameter variance was especially high for the weight CUSUM algorithm, which could explain the poor performances when compared to MACD.

Mean values for weight, impedance, and the respective output indices of the optimal algorithms during periods preceding a hospitalization compared to the other periods are shown in Table 4. A statistically significant difference was only found for the NITTI measurements and algorithms based upon NITTI.
### Table 3. Mean, standard deviation, and individual values for the estimated optimal parameters in each of the 8 folds created using the described stratified cross-validation procedure.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Body weight</th>
<th>Transthoracic impedance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV ^a step</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>RoT ^b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold</td>
<td>3.5 (0.08)</td>
<td>-0.31 (0.035)</td>
</tr>
<tr>
<td>Days</td>
<td>14.4 (3.7)</td>
<td>20.5 (1.41)</td>
</tr>
<tr>
<td>MACD ^c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold</td>
<td>0.8 (0.38)</td>
<td>-0.10 (0.014)</td>
</tr>
<tr>
<td>Short-term avg. window</td>
<td>8.6 (1.19)</td>
<td>8.1 (0.99)</td>
</tr>
<tr>
<td>Long-term avg. window</td>
<td>25.6 (10.84)</td>
<td>36.3 (3.54)</td>
</tr>
<tr>
<td>CUSUM ^d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold</td>
<td>11.0 (7.87)</td>
<td>-8.13 (2.65)</td>
</tr>
<tr>
<td>Days</td>
<td>26.9 (18.3)</td>
<td>18.8 (2.31)</td>
</tr>
<tr>
<td>Depreciation</td>
<td>1.13 (0.40)</td>
<td>0.75 (0.19)</td>
</tr>
</tbody>
</table>

^aCV: Cross-validation  
^bRoT: rule of thumb  
^cMACD: moving average convergence divergence  
^dCUSUM: cumulative sums

### Table 4. Population mean output index values for RoT, MACD, and CUSUM algorithms using the optimal parameters (see 2) in the 2-week period preceding a hospitalization compared to all other periods.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD) value in 2-week period before decompensation</th>
<th>Mean (SD) value in nond Decompensation periods</th>
<th>Statistical significance^d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (kg)</td>
<td>83 (10)</td>
<td>84 (19)</td>
<td>.97</td>
</tr>
<tr>
<td>Weight-RoT ^a (kg)</td>
<td>0.3 (1.2)</td>
<td>0.06 (0.87)</td>
<td>.76</td>
</tr>
<tr>
<td>Weight-MACD ^b (kg)</td>
<td>0.08 (0.30)</td>
<td>0.02 (0.22)</td>
<td>.24</td>
</tr>
<tr>
<td>Weight-CUSUM ^c (kg)</td>
<td>1.9 (2.7)</td>
<td>0.8 (1.3)</td>
<td>.58</td>
</tr>
<tr>
<td>TTI (log Ohm)</td>
<td>3.0 (0.3)</td>
<td>3.4 (0.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>TTI-RoT (log Ohm)^a</td>
<td>-0.07 (0.12)</td>
<td>0.00 (0.08)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>TTI-MACD (log Ohm)^a</td>
<td>-0.032 (0.044)</td>
<td>0.003 (0.028)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>TTI-CUSUM (log Ohm)^a</td>
<td>-6.4 (9.4)</td>
<td>-0.7 (2.0)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

^aRoT: rule of thumb  
^bMACD: moving average convergence divergence  
^cCUSUM: cumulative sums  
^dEstimated with a mixed-effect model with patient specific random effects. For the algorithms the cross-validation output was used.
Discussion

Principal Findings

The main finding of the present study is that change in NITTI is a stronger predictor of an impending decompensation compared to changes in weight (cross-validation estimate was 60% for NITTI-CUSUM vs 33% for Weight-MACD) and that both measurements benefit from trend detection algorithms. Mean values of NITTI in the 2-week period preceding a decompensation event were lower than in nondecompensation periods ($P<.001$).

Fluid overload is one of the leading causes for HF hospitalization and body weight increase has been linked to an increased risk of hospitalization [21]. However, directly applying a weight gain difference to predict imminent decompensation is challenging. This study corroborates the findings of Zhang [23] and Abraham [26], who also reported low predictive ability of alarms using short-term weight change. Short-term weight increase will detect a large and rapid fluid accumulation. Our evaluation of the rule suggested by the ESC guidelines is that it has high specificity but it is not a very sensitive method to predict HF hospitalization, as gradual weight increases are missed. A moving average algorithm focuses on progressive changes in weight, removing much of the inherent variability in weight measurements and errors due to the home setting in which patients might deviate from the measurement protocol, and daily changes due to dietary and fluid intake are averaged out. This could explain why lower threshold values led to higher sensitivity while still retaining specificity.

The increase in thoracic fluid due to congestion should decrease impedance measurements. Several studies have reported positive results from algorithms using impedance to detect decompensations [19,20,33]. To test algorithms proposed for decompensation detection using impedance measurements, we employed a cross-validation procedure to estimate performances. The results are similar, although on the lower side of what has been reported for ITI in terms of sensitivity (76.4% [26], 76.9% [33], 60% [34]), perhaps partly accounted for by the robust methods we employed. Reported performances from feasibility studies usually decline in later prospective studies [35], which the leave-subject-out protocol is designed to emulate.

Comparisons between predicted performances of weight and impedance measurements in Figure 3 show that impedance is the stronger predictor. This is also suggested by the analyses of the mean output index in the 2-week period preceding a decompensation (Table 4), for which a statistical difference was found compared to periods without decompensation for all impedance algorithms as well as the impedance value, but not for any of the weight algorithms. Abraham et al. [26] also showed a higher sensitivity for impedance measurements when compared to weight. However, we showed that the gap in performance could be made smaller with more sophisticated weight trend algorithms compared to the rules suggested by Abraham (in which the 3 rules with a specificity >90% had a maximum sensitivity of 7%). Sensitivity to fluid build-up in the lungs, whether through redistribution of fluids or retention, could explain the increased performance of impedance when compared to body weight [11]. Similarly, weight loss from malnutrition might mask fluid accumulation in weight measurements, which would still be picked up by NITTI. The focus in this study on high specificity algorithms might also have put weight algorithms at a slight disadvantage; evidence of this can be found in the stability analysis (Table 3), in which the high parameter variance for the weight-CUSUM algorithm could have resulted from the difficulty of finding a highly specific algorithm, which led to a negative impact on its cross-validated performance.

The difficulty in assessing prediction algorithms is known [36]. Different evaluation metrics can show diverging results, because they shed light on different aspects of performance. Definitions of what constitutes a true positive and false positive have a great effect on performance. In this study, we focused on algorithms with high specificity evaluated using 2-week intervals, with the best-performing alarm having a sensitivity rate of 60%. Although this catches several patients at a high specificity, it still raises unexplained alarms and has a relatively low positive predictive value of 10.9% for impedance and 3.2% for weight. A measure focusing on the workload associated with managing these alerts, such as false alarms per patient year has been used by several other studies as a surrogate specificity metric [26,33-35]. Defined as an alert not resulting in a hospitalization, the NITTI-CUSUM algorithm has a cross-validated estimate of 0.48 false alarms per patient year. These seemingly contradictory performance measures can be explained by the rarity of 2-week periods resulting in hospitalization, when compared to the full amount of telemonitoring data. An alarm that goes on for 5 weeks would cross three 2-week periods and could generate 3 false positives; however, using the false-alarm metric it would only add one false alarm.

Therefore, the positive predictive value of 10.9% should be seen in the context of 2-week windows having both high specificity and sensitivity and compared to the relatively low predictive value of current weight algorithms.

Low levels of positive predictive value have also been observed in many other studies evaluating prediction algorithms from daily measurements [35,37,38]. The concept of predicting future events might be less realistic than providing indications that could be acted upon. This approach could tailor actions depending on which monitored sign was detected. Indeed, many signs that have been linked to deterioration, for example, arrhythmias [39], breathing rates [38], and heart-rate variability [40], can be detected noninvasively and may be included in such an approach. Importantly, the implementation of better decompensation algorithms will reduce the number of clinical alerts that would need to be dealt with by a telehealth nurse or physician. This will result in better resource utilization, with the management of larger patient caseloads and, therefore, a reduction in the costs of patient management.

Limitations

Although clinicians were blinded to the observational data, they could have intervened based on increased weight data for worsening patients. If such interventions did not result in a hospitalization, they were not recorded in this study and might have negatively affected the results. In the SENSE-HF trial [37],
a substantial increase in positive predictive value was reported after including signs and symptoms of worsening HF diagnosed by a physician rather than only adjudicated HF hospitalizations; therefore, it could also be expected that several false positives were due to “mild” decompensations. Indeed, it is possible that patients often self-correct decompensation by reducing dietary salt, increasing adherence to medication, or even by taking extra doses of diuretic. Changes in environmental temperature might also affect compensation. In this study, high specificity alerts were explored. However, sacrificing specificity for improved sensitivity may be a good complement if management of alerts can be handled by patients without resorting to professional advice. Combining specific alerts with a strategy of health maintenance might be superior to one of only crisis detection and management [41]. Most patients are interested and able to contribute to their care if they are given the information and confidence to do so. Remote monitoring provides a safe environment or safety net to encourage such behavior.

Incorrectly using the measurement equipment could have caused erroneous values with the net effect of lowered performances. The surface on which the scales sit, their accuracy, clothing, and use by other family members can all cause problems with measurement. Bio-impedance weight scales (a different technology from NITTI) require patients to remove their socks and shoes and hence may improve the consistency of measurement. Giving patients feedback and asking them to recheck their weight if it falls out of the expected range are all likely to improve the data quality on which the algorithms are based. The limited amount of data available for this study makes generalizations difficult. Application of cross-validation procedures were employed to minimize this effect; however, the calculated percentage values were ultimately derived from a small set of subjects and should therefore be seen as qualitative indicators of performance.

Conclusion

Daily measurements of transthoracic impedance using a vest with textile electrodes is a feasible way to monitor HF and provides a more accurate indication of upcoming decompensations when compared to weight for all 3 algorithms tested (RoT, MACD, and CUSUM). Trend detection algorithms outperformed RoT measures suggesting that tracking the progression is more important than direct measures of change, which currently are suggested by guidelines. However, the low positive predictive value of all the algorithms tested did not allow accurate prediction of impending HF hospitalizations. Implementation of trend detection algorithms might better serve as indications of worsening, which, when integrated with other clinical measures, could be useful for treatment management. The promising results from this investigation warrant further trials with noninvasive TTI as a technology for the management of HF, perhaps connected to actionable alerts. These alerts would promote a strategy of “health maintenance” to keep the patient as close to their ideal state as possible on a daily basis, which could be combined with a strategy of “crisis detection and management” if the first strategy failed.

Acknowledgments

This work was supported by the EU Marie Curie Network iCareNet under grant number 264738. Data were provided by the MyHeart project, which was partially financed by the EU FP6 program under grant number 507816.

Conflicts of Interest

ICG is a PhD student employed at Philips Research. AGB, HR, and JH are employed by Philips Research. JGFC and KGM have received departmental research support from Philips.

Multimedia Appendix 1

Detailed description of algorithms to detect decompensated HF.

References


9. McMurray JJV, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur Heart J 2012 Jul;33(14):1787-1847 [FREE Full text] [doi: 10.1093/eurheartj/ehs104] [Medline: 22611136]


22. Whellan DJ, Ousdigian KT, Al-Khatib SM, Pu W, Sarkar S, Porter CB, PARTNERS Study Investigators. Combined heart failure device diagnostics identify patients at higher risk of subsequent heart failure hospitalizations: results from PARTNERS HF (Program to Access and Review Trending Information and Evaluate Correlation to Symptoms in Patients With Heart


**Abbreviations**

ACE: angiotensin converting enzyme  
ARB: angiotensin receptor blockers  
CABG: coronary artery bypass grafting  
CUSUM: cumulative sums  
CV: cross-validation  
ESC: European Society of Cardiology  
HF: Heart Failure  
HFSA: Heart Failure Association of America  
ICD/CRT: implantable cardioverter-defibrillator/cardiac resynchronization therapy  
ITI: intrathoracic impedance  
PAP: pulmonary arterial pressure  
PPV: positive predictive value  
NPV: negative predictive value  
PPV: positive predictive value  
ROC: receiver operator curve  
RoT: rule-of-thumb
Health Information Technology: Meaningful Use and Next Steps to Improving Electronic Facilitation of Medication Adherence

Abstract

Background: The use of health information technology (HIT) may improve medication adherence, but challenges for implementation remain.

Objective: The aim of this paper is to review the current state of HIT as it relates to medication adherence programs, acknowledge the potential barriers in light of current legislation, and provide recommendations to improve ongoing medication adherence strategies through the use of HIT.

Methods: We describe four potential HIT barriers that may impact interoperability and subsequent medication adherence. Legislation in the United States has incentivized the use of HIT to facilitate and enhance medication adherence. The Health Information Technology for Economic and Clinical Health (HITECH) was recently adopted and establishes federal standards for the so-called "meaningful use" of certified electronic health record (EHR) technology that can directly impact medication adherence.

Results: The four persistent HIT barriers to medication adherence include (1) underdevelopment of data reciprocity across clinical, community, and home settings, limiting the capture of data necessary for clinical care; (2) inconsistent data definitions and lack of harmonization of patient-focused data standards, making existing data difficult to use for patient-centered outcomes research; (3) inability to effectively use the national drug code information from the various electronic health record and claims datasets for adherence purposes; and (4) lack of data capture for medication management interventions, such as medication management therapy (MTM) in the EHR. Potential recommendations to address these issues are discussed.

Conclusion: To make meaningful, high quality data accessible, and subsequently improve medication adherence, these challenges will need to be addressed to fully reach the potential of HIT in impacting one of our largest public health issues.

(JMIR Med Inform 2016;4(1):e9) doi:10.2196/medinform.4326
KEYWORDS
medication adherence; compliance; health information technology

Introduction

Non-adherence to prescription medications is common and costly [1]. Approximately 20-30% of prescription medications are never filled [1], with approximately 40% of patients failing to fill an initial prescription [2-4]. Even after a medication has been acquired, many patients do not follow prescription instructions. Within one year, over 50% of patients prematurely discontinue their medications [2,5,6]. The problem of medication non-adherence is complex and pervasive with a lack of accountability dispersed across patients, their caregivers, clinicians, pharmacy benefits, and the health care systems as a whole.

While there is no universal solution to improve medication adherence, health information technology (HIT) can inform and accelerate ongoing strategies to initiate, improve, and monitor medication adherence. Studies increasingly demonstrate that the use of HIT can improve the quality and coordination of care and lead to better health outcomes [7,8]. Two commonly cited examples of HIT are electronic health record (EHR) systems and electronic prescribing (e-prescribing), the electronic generation of a medication prescription and its routing to a pharmacy.

Recent legislation in the United States has incentivized the use of HIT to address medication adherence [9]. The Health Information Technology for Economic and Clinical Health (HITECH) Act was passed in 2009 as part of the American Recovery and Reinvestment Act. It authorized an estimated US $30 billion in incentives to eligible professionals and hospitals to adopt and meet federal standards for the so-called "meaningful use" of certified EHR technology [9]. Meaningful use in the context of medications includes EHR functions such as e-prescribing, creating linkages between patient diagnosis and treatment plan, generating reports for clinical quality measures, integrating clinical decision support, and providing electronic messaging between providers. Electronic patient portals also allow for secure messaging between patients and providers, potentially improving medication-related communication and care while reducing the frequency of traditional in-office visits.

The Stage 2 core objectives include providing patients the ability to view online, download, and transmit their health information [10].

The Office of the National Coordinator recently proposed a version of the third stage of meaningful use, with a focus on both functionality and health care outcomes [11]. The Centers for Medicare and Medicaid Services (CMS) Stage 3 Meaningful Use proposed rule focuses on the use of EHR technology to promote improved patient outcomes and health information exchange. The rule proposes to ensure health systems and providers are coordinating care for patients, providing patients with easy access to their health information, and fostering data collection in a format that can be shared across multiple health care organizations. Early drafts of the Stage 3 Meaningful Use rule included medication adherence as well as several other related factors, such as medication reminders for refills. However, in early 2014, these were removed from the Stage 3 plan and are now included at the discretion of providers [12].

Meaningful Use Stage 3 would address key gaps identified in EHR functionality that impedes improvement in clinical outcome and are considered essential for adoption by all providers (Figure 1). These key gaps include information exchange among provider entities, connectivity for patient engagement, and technology advances to reduce disparities by providing decision support for national high-priority conditions. Each of these gaps has implications for medication use, including opportunity to improve monitoring, documentation, communication, and feedback [12]. Thus, these proposed changes for Stage 3 Meaningful Use should streamline previous meaningful use criterion while improving health care quality.

Despite the well-intentioned design of certification criteria and clinical quality metrics, the status of medication adherence in the meaningful use guidelines is unclear. The purpose of this paper is to review the current state of HIT as it relates to medication adherence programs, acknowledge the potential barriers in light of current US legislation, and provide recommendations to improve ongoing medication adherence strategies through the use of HIT.
Results

Challenges of HIT Interoperability, Connectedness, and Reciprocity

Interoperability, defined as the extent to which systems and devices can exchange and interpret shared data [13], has been a longstanding challenge for meaningful use initiatives related to medication adherence. Lack of data connectivity and reciprocity (i.e., the extent to which systems can interpret shared data, aka data exchange) across settings of care delivery has been identified as a key factor in poor data capture [14,15]. Specifically, interoperability that enables patients to share health information with their provider and/or health care system (e.g., smart pill storage containers or blister packages that record and transmit when a medication has been taken and communicates with an EHR) in an actionable form is underdeveloped. As a result, patients are not able to "connect" data from self-monitoring efforts to their health records and providers are not able to evaluate these data in the context of the person’s other health information. In the case of medication adherence, there are a number of "users" in addition to the patient and provider. These include pharmacies, health systems, and payers. This network of "users" may become extensive; for example, patients may "price shop" and use multiple pharmacies to fill their prescriptions, intensifying the need to share information. As these stakeholders seek to evaluate and improve medication packaging and delivery devices the number of users in the network expands. In addition, their respective motives, goals, and willingness to participate in a transparent and interoperable data platform becomes less predictable. The end-result of a successful, "interoperable exchange" must be representation of data in a user-accessible format that allows users access to these data.
The problem of poor medication adherence HIT data reciprocity across systems takes the following two forms: (1) the complexity of the data, coding syntax, and the transmission infrastructure on which these data reside, and (2) the output, such as pill-taking history, medication refill rates or patient-reported experience of side effects. In both cases, these data are not accessible, timely, nor easily understood by the health care system. Thus, providers, patients, family members, and community support often lack adequate communication regarding medication use. The first gap, underdevelopment of data reciprocity across clinical and home settings, gives rise to the second problem of poor adherence outcome. Ironically, the challenge of accessibility and usability is exacerbated by innumerable handheld device apps that are accessible, timely, and easily understood, and are increasing in number. As of October 2013, for example, 160 apps were identified as being available and focusing on an aspect of medication adherence. The sheer number of available apps makes it difficult for patients and providers alike to identify the best solution to address their unique needs. Moreover, creating a system to support sharing of data between this multitude of apps and an array of EHR systems is daunting. Given that medication management requires many participants, roles, responsibilities, and handoffs, data capture that is meaningful for patients is often lost in the chaos of medication management as care processes cross boundaries and settings of care. Thus, criteria for evaluating and encouraging medication data reciprocity are warranted.

An example of underdevelopment of data reciprocity is demonstrated in the CMS-funded Southeastern Diabetes Initiative (SEDI) program [16]. This ongoing project consists of behavioral strategies to support medication-taking that are broadly implemented through clinics, community venues, and patient home visits in four southeastern US counties in North Carolina, Mississippi, and West Virginia. The EHRs in each of the four county sites lack designated fields for documenting patient participation in medication management and self-care support interventions. This gap renders the task of measuring and collecting feedback regarding the impact of medication related care processes impossible, and key stakeholders such as community health workers, nurses, dieticians, physicians, and pharmacists lack data to support communication and provide feedback to patients regarding participation in adherence interventions.

A solution, one which obviates the need to directly link disparate health information systems containing sensitive protected health information (PHI), lies within the range of possibilities presented by Meaningful Use Stage 3. Such a solution would require functionality in the form of common discrete identifiers for EHR fields to indicate patient-selected medication management resources and the discrete identifiers for community-based resources that were used to access, deliver or monitor the management strategy. For example, participation in the Diabetes Self-Management Program [17,18] would be documented independently by county providers in the EHR, but the patients’ enrollment and actual attendance in such programs would be in a sortable, discrete, commonly labeled data field in the EHR, regardless of county. Subsequent changes in glycated hemoglobin (HbA1c) could be easily identified and evaluated, both at the individual and aggregate levels, and would be available immediately. In this way, EHR functionality is linked with medication management outcomes across settings, from inpatient to outpatient, and across systems of care from county to county.

Improving linkages and data access will likely improve clinical care. Linkages provide an effective method to evaluate quality indicators related to medication adherence including patient-centered outcomes (PCOs) that are associated with improved medication adherence. Better integration and capture of PCOs will become increasingly important as they are integral components of a more comprehensive approach to patient care, and include objective measures of medication-related clinical outcomes (eg, blood sugar, glycated hemoglobin, blood pressure, and lipids). In addition, the patient-reported outcomes associated with medication management including knowledge, side effects, beliefs of medications are important subjective indicators of patient engagement, and progress in medication-related goals. Patient-reported outcomes (eg, cost-motivated medication non-adherence, barriers to adherence) assessed via validated surveys and questionnaires are currently in use in many clinical settings, but are inconsistently represented in available electronic formats. As a result, patient-reported outcomes are inaccessible for research and unable to be monitored for data quality. These patient-reported outcomes include common measures such as the (hospital) Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys, health literacy measures, and medication adherence measures such as the Morisky Medication Adherence Scale [19] or the Medication Discrepancy Tool [20,21]. Improved linkages and data access will therefore improve the use in clinical care of patient reported outcomes data associated with medication-taking that are collected across healthcare systems and clinical data networks.

In summary, the state of the science of HIT falls short of this definition of interoperability. Though the phases of meaningful use implementation have pushed increasing numbers of hospitals including critical access hospitals and federally qualified health centers, to invest in EHRs, these systems are challenged by weak links in the data definitions across electronic platforms and low levels of interpretability and access by both physicians and patients [22,23]. Thus, the potential beneficial impact of HIT on medication adherence has not been achieved.

Solutions will require functionality in the form of discrete identifiers that are commonly defined and consistently adopted and applied at each point of data contact. These points of contacts will range from data capture, to coding, to data transmission, and include outpatient data such as pill-taking history, medication refill rates, and patient-reported experience of side effects. Each of these data components must be accessible, timely, and easily understood by users.

**Inconsistent Data Definitions**

Across almost every level from descriptive evaluation of medication fill-rate patterns and trends [24] to complex predictive modeling of the association of medication adherence with clinical outcomes [25,26], methods for analysis of medication adherence fail to meet patients’ needs and expectations. This relative lack of progress is due in part to
issues surrounding the data itself. The variability in data definitions and inconsistency in terms used in practice and research prevent successful application of interventions from controlled research settings to real-world populations [27,28]. "Data definition" is the electronic specification established for keyed data entry of each data element. For example, blood pressure must be specified by type (arterial systolic, diastolic or mean), source (cuff, intra-arterial or venous), and limits on digital display (eg, two decimal places). Implications of underdeveloped data linkages are apparent when examining current methods for the evaluation of interventions to improve medication use. In addition, the definitions and terms most commonly used to evaluate medication use in research have not been conceptualized from a patient perspective [29,30]. Lastly, measurement and data capture from the many entities that contribute to management of adherence over time, including patients, caregivers, providers, communities, and health systems, are inaccessible or absent [31]. Gaps exist in the validity, accessibility, and efficiency of data sources commonly used to study adherence interventions and improve patient management of medications.

An example of inconsistent data definitions for medication adherence is the distinction between initiation, renewal, and therapeutic discontinuation. For many HIT systems, there is no mechanism in the health IT infrastructure that allows a prescriber to note that they are initiating a prescription for the first time (new to therapy) versus a refill authorization or new prescription for therapeutic continuation. While there is an existing mechanism for a prescriber to declare that they are discontinuing a medication in the 10.6 SCRIPT e-prescribing standard, for example, vendors have been slow to adopt this feature due to lack of a clear directive through meaningful use standards to date. Clinical programs such as medication reconciliation and pay for performance (eg, Medicare Star Ratings) could undergo dramatic improvements in precision through increased development, standardization, and incentive programs for adoption of order entry capabilities that have explicit data on medication initiation and discontinuation.

The underdevelopment of data linkages has major impediment for improving health care quality and patient outcomes. An outdated or incomplete medication list may give health care providers insufficient or inaccurate information to provide proper medication management of patients’ conditions. Similarly, there may be insufficient or inaccurate data to understand a patient’s non-adherence. Without incentives for medication reconciliation, this problem may continue unchecked. By not having consistent data definitions for non-adherence, it is unclear what would be "flagged" and trigger intervention from the health care system. Is non-adherence filling a prescription seven days too late or not at all? If a patient fails to fill a new prescription, does that constitute non-adherence? Having consistent data definitions would improve understanding of a patient’s ongoing medication use, which would in turn have the potential to improve symptom control and reduce morbidity and mortality.

The inconsistent data definitions and lack of harmonized data standards make tracking data relevant from a patient’s perspective difficult. These data sources include a patient’s perceptions of the medication, barriers and/or facilitators to taking the medication, and potential side effects; all factors likely to impact medication adherence. These data sources lack harmonized data definitions and data linkages across settings, resulting in poor accessibility of data from outpatient, community, and home settings to hospital and pharmacy dispensing industry settings. As a result, data are difficult to use, and when they are accessible, the data quality is poor and not amenable for use in research.

In addition to having inconsistent data definitions, there is also lack of agreement regarding where data should be stored; drug therapy problems typically do not have a place to reside within EHRs. Drug therapy problems are typically not included in the more general problem list within EHRs and other electronic systems. Many commonly understood drug therapy problems are the direct progenitor for the patient non-adherence and are critical pieces of information that are neither stored nor shared across HIT systems.

Challenges in identifying common data definitions that are meaningful to patients stem from the etymological origin of the terms used in research and practice to reflect "medication-taking". These terms were not derived from a patient perspective. Thus, terms that are most commonly used and most likely to be well defined and standardized [32], such as medication possession ratio or proportion of days covered, do not reflect aspects of medication-taking that are considered important or useful to patients. As a result, study questions and adherence interventions are less likely to be designed from a perspective that will yield meaningful information for patients. Thus, solutions need to include terms, data definitions, and data standards that are clearly defined, standardized, and oriented from a patient-centric view of medication management in everyday life.

Inability to Effectively use the National Drug Codes

A third major barrier to the use of HIT for medication adherence in the United States is the inability to effectively use the National Drug Code (NDC) information (US Food and Drug Administration (FDA), 2013) from the various EHR and claims datasets for adherence purposes. Though NDC codification provides a unique 10-digit drug identifier with three segments for each FDA approved medication, only the first segment is a fixed identifier from the FDA, while the second and third segments vary by company and product. As a result, although NDC codification is intended to greatly improve data quality, the 3-segment numeric indicator of the vendor, product specification (strength, dose, and formulation of the drug), and trade packages are not standardized. Thus, the stage at which the codes are implemented in the EHR system is critical [33]. One solution to address the inability to effectively use the NDCs could be improved through standardized codes across all contexts that are integrated into the EHR, much like existing International Classification of Diseases (ICD) codes.

Capture of Medication Management Therapy

A fourth challenge with HIT in regards to medication adherence is the poor capture of medication management therapy (MTM)
in current electronic data systems. MTM ensures optimum therapeutic outcomes through improved medication adherence, reduces the risk of adverse events, is developed in cooperation with licensed and practicing pharmacists and physicians, and is coordinated with any care management plan established for a targeted individual under a chronic care improvement program [34]. Specific activities of MTM include performing a comprehensive medication review, formulating a treatment plan, and providing patient education to promote adherence, among other services. Although MTM is part of Medicare Part D, data regarding provision of MTM services is not routinely captured and is not available (eg, "traceable") beyond the initiation of the checkbox at the pharmacy. Functionality "fix" to address this gap would be to link MTM data from pharmacy databases to the EHR such that actionable alerts could be routed to the physician message basket or "inbox", and summaries of educational interventions would be logged in the outpatient interactions log and tagged with "pharmacy provider" label for rapid identification. In this way HIT connectivity closes the communication loop and meets guideline measures for communication and information exchange as outlined in meaningful use stage three.

Some of the data elements that are most important to patients are not captured at all. For example, the data elements reflecting daily management of medications by patients, such as the implementation of routines and reminders to facilitate medication-taking, the monitoring of daily physiologic indicators that drive dose, such as blood sugar in diabetes or daily weight in heart failure, or the access to pharmacies with home delivery; these details of daily management significantly alter the ability of patients to manage medicines. Yet, most are not captured, or if captured are not transmitted to providers in a way that enables feedback using real-time data. The Medicare Part D Medication Therapy Management program is limited to measures of cost and resource utilization. Though one study has reported improvements in patient outcomes [35], these types of intervention outcomes for management of medicines are not consistently documented and addressed.

Thus, one solution to the poor capture of MTM in current electronic records would be to link MTM data from pharmacy databases to the EHR such that actionable alerts could be routed to the provider message basket or "inbox". These messages could summarize outpatient interactions and be tagged with "pharmacy provider" label for rapid identification.

Role of Mobile Apps, HIT, and Medication Adherence

In addition to the challenges laid out above, consideration of how mobile medical apps will play into the mix of operability, HIT, and medication adherence are needed. The US FDA issued final guidance for developers of mobile medical apps, which are software programs that run on mobile wireless communication devices and perform the same functions as traditional medical devices. The FDA intends to focus its regulatory oversight on a subset of mobile medical apps that present a greater risk to patients if they do not work as intended. For example, an app that allows a health care professional to make a specific diagnosis by viewing a medical image from a picture archiving and communication system on a mobile phone or tablet. Another example requiring FDA regulation would be transforming a mobile platform into a regulated medical device such as an app that turns a mobile phone into an electrocardiography machine to detect abnormal heart rhythms or determine if a patient is experiencing a heart attack. However, further consideration of how mobile apps can interact with electronic medical records as well as provide more information than simply reminders is needed. For example, consider a mobile app is able to collect information from the patient regarding chemotherapy side effects they may be experiencing, then uses an evidence-based algorithm to prioritize these side effects, and integrates the prioritized list into the EHR for the provider to evaluate.

Discussion

A system-based view of overall medication use, management, and patient adherence is needed. Improving medication use is a systems challenge, given the many entities involved in the whole process. Patients reliant on medications are tied to the prescriber’s office, the dispensing pharmacy, their home, their health plan, prescription drug plan, and pharmacy benefit management. A system-based view of medication use would be a first step towards building a model that would allow stakeholders to track a patient’s experience with medications over an entire continuum of care, and allow stakeholders to visualize how and when different interventions are (or should be) delivered to patients over time.

Until the challenges of data complexity are addressed, health care providers may be reticent to have medication use data incorporated into the EHR, unless there is clear guidance on how to respond to various lapses in medication use, especially as provided for in real-time. There may be additional medical liabilities that might occur if providers fail to respond to a prompt indicating non-adherence, which results in a negative outcome for the patient. Currently, liability for non-adherence lies primarily with the patient. Integrating these data into the EHR, as proposed, will likely shift liability to the provider and require that expectations and standards for response roles and timing be considered.

Finally, to bring a system-based perspective to bear on complex data and data integration issues, we summarize six recommendations for solutions to move HIT to the next phase of use (Textbox 1). These six recommendations are not an exhaustive list, but encapsulate the evidence to date and the opportunities that lie ahead. Importantly, these potential solutions present an opportunity for collaborative work across key stakeholders, including patients, providers, pharmacists, payers, and health technology programmers, designers, and developers. Building broad, collaborative, and multidisciplinary working groups to address these issues is the next exciting frontier in health.

http://medinform.jmir.org/2016/1/e9/
Textbox 1. Recommendations to leverage HIT to improve medication use and adherence.

<table>
<thead>
<tr>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Improve HIT interoperability by designating common discrete fields reflecting medication management.</td>
</tr>
<tr>
<td>2. Develop consistent data definitions for medication management activities and outcomes.</td>
</tr>
<tr>
<td>3. Increase use of national drug code by using these codes for prescribing and dispensing.</td>
</tr>
<tr>
<td>4. Develop data capture in EHR for MTM and allow the sharing of MTM strategies to providers across the system. In addition, having the ability to acknowledge that further clarification regarding medication if need be and allowing the sharing of this information to providers in the patients’ health care network.</td>
</tr>
<tr>
<td>5. Leverage integration of mobile apps to improve patient self-monitoring data capture and associated provider feedback when appropriate.</td>
</tr>
<tr>
<td>6. Capture the data linkages for medication adherence from the provider perspective that highlights therapeutic initiation, continuation, and discontinuation.</td>
</tr>
</tbody>
</table>

Conflicts of Interest

Dr Bosworth has received funds from Sanofi, Takeda, MeadWestVaco, Improved Patient Outcomes, and Johnson and Johnson to Duke. He also reports funds for consulting from Walgreens, CVS/Caremark, Blue Cross/Blue Shield of Arkansas, Genentech, Sanofi, Takeda. Dr Phil Mendys is employed by Pfizer. Dr Christopher Granger has received funds from Bristol Myers Squibb, GSK, Boehringer Ingelheim, Medtronic Foundation, Pfizer, Sanofi, Takeda, Astra Zeneca, Daiichi Sankyo, Bayer, Astra Zeneca, Janssen, and Armeotheron. He also reports funds for consulting from Boehringer Ingelheim, Bristol Myers, GSK, Hoffmann-La Roche, Eli Lilly, Pfizer, Sanofi, Takeda, The Medicines Company, Astra Zeneca, Daiichi Sankyo, Janssen, Bayer, Gilead, and Medtronic Inc.

References

12. The Office of the National Coordinator for Health Information Technology. Moving Forward Towards an Interoperable Learning Health System: Improving Flexibility, Simplicity, Interoperability and Outcomes to Achieve a Better, Smarter
and Healthier System. 2015. URL: https://www.healthit.gov/sites/default/files/CMS-Stage-3-Meaningful-Use-proposed-rule%20FactSheet.pdf [accessed 2015-10-27] [WebCite Cache ID 6cb02THZI]


Abbreviations

CMS: Centers for Medicare and Medicaid Services
EHR: electronic health record
FDA: Food and Drug Administration
HIT: health information technology
MTM: medication management therapy
NDC: National Drug Code
PCO: patient-centered outcome

©Hayden B Bosworth, Leah L Zullig, Phil Mendys, Michael Ho, Troy Trygstad, Christopher Granger, Megan M Oakes, Bradi B Granger. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 15.03.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.

http://medinform.jmir.org/2016/1/e9/
Conceptual Models in Health Informatics Research: A Literature Review and Suggestions for Development

Kathleen Gray1*, PhD; Paulina Sockolow2*, DrPH

1Health and Biomedical Informatics Centre, University of Melbourne, Melbourne, Australia
2Health Systems & Sciences Research, College of Nursing and Health Professions, Drexel University, Philadelphia, PA, United States
*all authors contributed equally

Corresponding Author:
Kathleen Gray, PhD
Health and Biomedical Informatics Centre
University of Melbourne
Level 1
202 Berkeley Street
Melbourne, 3010
Australia
Phone: 61 3 8344 8936
Fax: 61 3 8344 8936
Email: kgray@unimelb.edu.au

Abstract

Background: Contributing to health informatics research means using conceptual models that are integrative and explain the research in terms of the two broad domains of health science and information science. However, it can be hard for novice health informatics researchers to find exemplars and guidelines in working with integrative conceptual models.

Objectives: The aim of this paper is to support the use of integrative conceptual models in research on information and communication technologies in the health sector, and to encourage discussion of these conceptual models in scholarly forums.

Methods: A two-part method was used to summarize and structure ideas about how to work effectively with conceptual models in health informatics research that included (1) a selective review and summary of the literature of conceptual models; and (2) the construction of a step-by-step approach to developing a conceptual model.

Results: The seven-step methodology for developing conceptual models in health informatics research explained in this paper involves (1) acknowledging the limitations of health science and information science conceptual models; (2) giving a rationale for one’s choice of integrative conceptual model; (3) explicating a conceptual model verbally and graphically; (4) seeking feedback about the conceptual model from stakeholders in both the health science and information science domains; (5) aligning a conceptual model with an appropriate research plan; (6) adapting a conceptual model in response to new knowledge over time; and (7) disseminating conceptual models in scholarly and scientific forums.

Conclusions: Making explicit the conceptual model that underpins a health informatics research project can contribute to increasing the number of well-formed and strongly grounded health informatics research projects. This explication has distinct benefits for researchers in training, research teams, and researchers and practitioners in information, health, and other disciplines.

(JMIR Med Inform 2016;4(1):e7) doi:10.2196/medinform.5021

KEYWORDS
medical informatics; theoretical models; conceptual framework; conceptual model; design-based research; implementation research; evaluation research; health informatics; research design; research training

Introduction

Conceptualizing Research in Health Informatics

There is consensus that the discipline of health informatics is characterized by the integration of elements from many other fields of knowledge. The components of health informatics, apart from the biomedical sciences, include computer science, information science, decision science, statistics, cognitive science, organizational theory, and others [1]. In essence, health...
informatics is “cross-training” between broadly defined information sciences and health sciences [2]. Ideally, research rests on “methodologies that capture the processes integral to applications, the users and the world in which the users function” [3]. However, a growing number of researchers who do not identify themselves as health informaticians are now doing research into the design, implementation, and evaluation of information and communication technologies in the health sector. This growth is fuelled by new technologies that reduce the health professionals’ barriers to application development, and also by the growing market for consumer technologies that are not subject to medical device regulations. For example, the current emergence of apps for mobile phones and the increased ease of programming these apps is said to “enable busy clinicians to develop simple mobile Web-based apps for academic, educational, and research purposes, without any prior knowledge of programming” [4]. The ensuing research appears in the journal and conference literature of a variety of fields including clinical specialties, health policy, information management, and software engineering, to name a few. This paper is aimed at researchers still in training, or practitioners new to the field who wish to align their work more strongly with the discipline of health informatics.

Taking a disciplined approach to health informatics research means operating across the component domains of expertise by using integrative conceptual models. A working definition of a conceptual model is that it is an explanation of the researcher’s thinking about the key constituents of the research problem, and why the whole problem is greater than the sum of its parts because of the way these interconnect and interact. In any field of knowledge, using a conceptual model to describe something about a subset or an aspect of the domain has value; that is, a conceptual model makes explicit the intended meaning of terms and concepts used and avoids ambiguity and misinterpretation. The terms conceptual framework and conceptual model are used interchangeably in the literature [5], and we use the latter throughout this paper. Health informatics conceptual models connect the knowledge and thus explain the research in the language of two broad domains, health science and information science, can help to ensure that research is effective and has impact.

Too often, research on information and communication technologies in the health sector appears to miss either the health problem or the information technology problem. Some examples of missing the health problem are: a review of 55 heart failure risk computational models noted that few had been implemented in clinical practice [6]; a description of a technical solution to a perceived clinical problem omitted any mention of consultation with clinical experts [7]; and a description of a technical solution to a clinical issue did not fit the clinical workflow [8]. If the research does not capture the processes integral to both the world of health science and the world of information science, valuable efforts are expended developing applications that do not address the intended problem. Many innovations fail to achieve sustainability or other measures of success because “the current development of eHealth technology often disregards the interdependencies between technology, human characteristics, and the socioeconomic environment, resulting in technology that has a low impact in health care practices” [9]. Evaluation research too may fall short of offering key insights. For example, researchers who evaluated an electronic health record implementation using the Delone and McLean framework for evaluating generic information systems success recognized that they would have done better to develop a more health-specific approach to evaluation [10], and indeed other researchers had customized Delone and McLean for evaluating health-specific information systems [11].

Recognizing Conceptual Models in Health Informatics Research

Researchers new to the field of health informatics need to learn how to work with its conceptual models. However, this may not be taught formally, and exemplars and guidance in the literature are sporadic and scattered. Possibly due to publication word limits, much of the published health informatics research conveys an absence of discussion, even a lack of awareness, about the importance of conceptual models. In addition, papers may mention a conceptual model approach without specifying how it came into existence.

Some descriptions of the development of conceptual models in specific health informatics research studies are available. Gordon et al described how models of clinical guideline knowledge had to be integrated with models of health care activities and processes in a conceptual model approach for automating distribution of clinical guidelines [12], and Ruland and Bakken enumerated the components of a conceptual model to support inclusion of patient preferences in clinical decision making and underscored the importance of incorporating knowledge from four domains [13]. In addition, Kaplan and Shaw compared a variety of ways that evaluation researchers have conceptualized the complex social and institutional dynamics of health information technology implementations [14], and Yusof et al [15] explained how theories from information science and evidence from health science informed the structure of a human-, organization-, and technology-fit evaluation framework for health information systems. There is no single right or wrong conceptual model that brings order to a set of ideas about a health informatics problem. The exercise of making the conceptual model explicit in a study, in words and/or figures, is critical to clarify what is known and to identify what is in question or not known, from the perspective of the researchers. Through the process of debating the merits of alternative conceptual models, their explanatory power, completeness, and other aspects about how well they represent the research objectives, new theories are formed.

This paper aims to promote more explicit use of health informatics conceptual models in research on information and communication technologies in the health sector, and encourage discussion of these conceptual models in scholarly forums. Learning how to develop and apply integrative conceptual models is an educational issue for researchers in training, and so too for those who train them and those who review their work. Using conceptual models has research significance for building the discipline of health informatics and benefits for many stakeholders in health informatics research.
Methods

A two-part method was used to structure and illustrate ideas of how to work effectively with conceptual models in health informatics research. First, we conducted a search of the literature of conceptual models, and then we used a qualitative research process to formulate a step-by-step approach to developing a conceptual model.

We looked for papers published up to 2014 that described the development of conceptual models in health information science and technology research in PubMed, IEEE Xplore, ACM Digital Library, Scopus, and Web of Knowledge. First the Medical Subject Heading (MeSH) "models, theoretical" paired with "medical informatics" were used. Then search terms were widened to include "research design" (especially where there was discussion of why a design was chosen). An additional search looked for possible pairings of "design science" or "design-based research", "implementation", and "evaluation" with health information systems and technology and with electronic health (e-health). Selection of a cross-section of full papers that made substantial mention of conceptual models was based on reading abstracts and also on mining reference lists from selected papers for further examples.

Critical reflection was used to formulate a step-by-step approach to developing a conceptual model. We examined the assumptions embedded in our experiences, associated them with a range of different factors, re-evaluated them using external reference points, and re-worked our ideas and practices [16]. Specifically, we drew on our separate experiences working in multidisciplinary health information technology research teams internationally over five years. We reconsidered the bases of our expertise as researchers and also as reviewers, supervisors, advisors, and examiners of research in health informatics. We analyzed the literature we had retrieved, looking at ways authors named, explicited, and sequenced key components in the development of a conceptual model. We agreed on specific steps in development of a conceptual model and went back to the literature and to our experience repeatedly for examples.

Using this method, we produced a set of suggestions, which remains untested in terms of its technical validity and sufficiency. Nevertheless, after formulating these suggestions, we found that teaching novice researchers to use conceptual models in a step-wise manner is recognized as effective by academics in other fields [17]. We also found that a similar idea had appeared in the literature of a different field of health research, namely health program evaluation [18]. The existence of these peer-reviewed publications added external validity to our method.

Results

The process of critical review and reflection led us to a consensus that supported a seven-step approach to developing a conceptual model for health informatics research. In this section we offer these steps and examples from the literature as a guide to the novice health informatics researcher. The methodology for working with conceptual models in health informatics involves (1) acknowledging the limitations of health science and information science conceptual models; (2) giving a rationale for one’s choice of an integrative conceptual model; (3) explicating a conceptual model verbally and graphically; (4) seeking feedback about a conceptual model from stakeholders in both the health science and information science domains; (5) aligning a conceptual model with an appropriate research plan; (6) adapting a conceptual model in response to new knowledge over time; and (7) disseminating conceptual models in scholarly and scientific forums.

Acknowledge the Conceptual Models of Contributing Domains

It is important to acknowledge that both health science and information science use a variety of conceptual models to represent entities and relationships in their respective domains; however, these fields of knowledge are differentiated by their approaches to conceptualizing problems. A defining characteristic of any given health informatics research problem is how it responds to the challenge to acknowledge the applicability and also the inadequacy of conceptual models from both health science and information science. That is, you should be able to describe some parts of the research question using concepts that appropriately represent the problem to the separate audiences or stakeholders from each domain. You must also go beyond this, to frame the problem in a conceptual model that transcends and integrates these domains.

In the information sciences, a conceptual model may also be referred to as a domain model. Furthermore, conceptual modeling should not be confused with other modeling disciplines, such as data modeling, logic modeling, or physical modeling. Many different conceptual models may be used for demonstration, optimization, construction, simulation and other activities in the application domain [19]. In the field of information science, recent examples in the literature can be found that discuss specific conceptual models in detail [20-22]. The complexity of health is a major reason why health informatics is not just another application domain in information science [23]. In the health sciences, an introduction to a range of conceptual models for defining and conceptualizing health argues that simplistic definitions of health lead to equally simplistic measures of health, health outcomes, and quality of care [24]. In the field of health, recent examples which explore conceptual models can be found [25-28].

Review Conceptual Models Already Used in Health Informatics

The next step is to review the health informatics literature for conceptual models that can be either applied or adapted to the research question. A clearly described search strategy for reviewing this literature is an indicator of the rigor you need to apply to the process of conceptualizing the problem [29,30]. The starting point is to compare various overarching conceptual models based on the power each may have to fully account for all the elements of the problem as you have chosen to define it. The introduction section of this paper offered examples from the 1997-2008 decade, and more recent examples can be found in references [9,31-34]. It is essential to justify your choice of a pre-existing conceptual model as it relates to the key features

http://medinform.jmir.org/2016/1/e7/
of your research problem. Alternatively, you may conclude that no pre-existing, cross-cutting conceptual models are adequate to explore this problem. This judgment also requires justification, and it opens the way for thought experiments about options for combining and elaborating the particular conceptual models that you have previously acknowledged. In one example, an investigation of a novel and under-researched technology in health care was able to proceed by integrating concepts of evidence-based treatment (from health science) and technology affordances (from information science) into a new conceptual model of therapeutic affordances of social media [35].

Schematize the Chosen Conceptual Model of the Research Problem

Making a schematic representation of your chosen conceptual model captures and refines the thought processes behind your choice. A visual artifact in the form of a diagram, motif, map or other type of figure (eg, a foil or straw man) can be used for reacting to and testing the thinking about a problem, to guide collaboration, and to assess research progress and outcomes. An example of schematizing a health informatics conceptual model to represent the relationships among health information technology characteristics is the Health Information Technology Reference-based Evaluation Framework [36]. A second example illustrates the temporal dimensions of five measures of health information system adoption in the Clinical Adoption Meta Model [37].

The challenge for every health informatics researcher is to think deeply about an apt way to visualize the specific problem space. Part of the contribution that your research makes to the field is determined by the originality you show in this step. Questions to consider include: How does the visualization of your conceptual model position the information science and the health science elements of the problem (eg, side-by-side versus above and below)? Does the level of detail match the intended level of investigation (eg, evaluating the impact of a policy may need to represent issues at a macro level or assessing software functionality may need a finer grained picture)? Does it leave too much to be inferred (eg, not indicating the direction in which a multi-part image should be "read") or use conventions in an unconventional way (eg, using the colors red for "go" and green for "stop")?

The schema also needs to be interpreted in words that explain it to someone who is unfamiliar with the research problem or who does not have access to the graphic. It should be clear why you have chosen the visual representations you are using to represent the key entities and relationships that your problem involves. For example, a study of unfulfilled and unrecognized or hidden health information needs explained the research framework using the graphic image and written analogy of an iceberg [38].

Your first attempt to represent the entities and relationships in your research problem using a Venn diagram, matrix, or flowchart may not suffice to give adequate detail or insight into the problem space. The visualization of knowledge is a field of study in its own right, and you may find it helpful to consult general works [39,40]. Although the visualization of data has become an active health informatics research area [41,42], the visualization of concepts is a very different order of activity.

Seek Critical Feedback on the Conceptual Model From Multiple Perspectives

Your conceptual model may appear completely sensible from your point of view. However, at this point in development, you should be thinking of your conceptual model as a communication tool. This tool should help you engage other people who are direct and indirect stakeholders in your research. Thus, stakeholders with other perspectives need to test its communicative power to assure you that it is making sense of the problem.

A conceptual model in health informatics research should pass the "goodness of fit" test for domain experts in both information science and health science. This is an informal but important step where you seek critique of your conceptual model from others who are at a distance from your research question. Their reactions allow you to refine and strengthen the supporting arguments for your conceptual model as needed. You are looking for toughness; now is the time to establish whether your conceptual model will stand up to scrutiny and be persuasive in a room full of either clinical specialists or computer scientists.

For a researcher in training, often the best way to ensure access to this kind of feedback is to make sure that the supervisory or advisory committee comprises people who bring information science and health science perspectives. Alternatively, the researcher needs to find suitable critical colleagues by tapping into networks of clinicians and researchers within the health and biomedicine community, and into networks of industry experts and researchers within the information science and technology community. Organizational mentoring programs, professional associations, or scientific societies can provide access to networks appropriate to the study.

If the research involves human participants, another method of seeking feedback on a conceptual model is to engage actively with prospective participants including patients and consumers. For example, Belanger et al. advocated involving patients from the inception of a research project centered on electronic health records [43].

Allow the Conceptual Model to Influence the Research Design

There must be a close connection between your conceptual model and research design. In a different field of health research, the alignment was expressed as follows:

The CF (conceptual framework) provided the basis for decisions about the development of a mixed-method research design and data collection measures. For each construct of interest, we determined the most suitable approach to collecting information. We arranged the constructs into two categories: those where validated quantitative measures were available (…); and second, those that were best suited for exploratory/qualitative methods. [18]
This is a good example of allowing the conceptual model to influence the research design, including the selection of the research procedures and outcome measures (for quantitative research) or themes (for qualitative research). Feasibility factors also influence research design, for example, access to sites for field studies, available funding and human resources, the exigencies of ethics approval, and time constraints. However, if any of these factors undermine the conceptual model to any great extent, then the research question and conceptual model need to be reformulated.

**Revisit the Conceptual Model in Light of the Research Findings**

Your conceptual model is what guides the approach you take to explore a real-world research problem. The corollary is that through exploring this problem, inevitably you are testing your conceptual model for its usefulness. While you are analyzing and discussing findings from your research project, you need to reflect on whether and how these findings support your initial conceptual model, identify where the conceptual model may need to be modified, and consider broader circumstances where the conceptual model may prove useful. This is an iterative and experiential process that stretches your thinking as your conceptual model "evolves and develops until it becomes refined and burnished, to emerge as a robust outcome of the research" [44].

In health informatics research, the gap between health sciences and information sciences that must be bridged by this step is substantial.

*In more mature fields such as medicine, it is standard practice, even mandatory, to conduct empirical research to evaluate the efficacy of proposed new practices prior to advocating their use (...). However in IS design research, it is often sufficient for researchers to argue on logical or theoretical grounds that their approach is effective.* [45]

How your health informatics conceptual model is tested depends on the research design appropriate to your study. In one example, thematic analysis of interview data is the basis for revising an initial conceptual model of health professionals’ mobile health use [46]. Examples where structured survey methods are the basis for revisions to initial conceptual models are found in studies of clinical information system success [47], and open access publishing use [48].

**Disseminate the Conceptual Model**

The final step is to disseminate your conceptual model through formal presentation and publication in scholarly and scientific conferences and journals. It is important to give over the time and space to include the conceptual model among the publications that come out of your research, so that it is captured and available for others to refer to and build on. You can structure a paper or presentation about your conceptual model using steps one through six in this paper; or examples cited throughout this paper offer many other successful models for publishing a description of your conceptual model.

When you describe your work on a conceptual model you make a contribution to the theory of health informatics, because your conceptual model adds a new theoretical representation of the entities and relationships in a problem space. In disseminating your conceptual model you address an acknowledged need in the field. The absence of theory in health informatics diminishes its status as a field of knowledge [49]; whereas, enunciating your conceptual model advances the field [3].

There are many diverse forums where it is possible to disseminate your conceptual model. This is one advantage of the diffusion of published research in health informatics. Apart from those already mentioned in this paper, three further examples of particular health informatics conceptual models published in recent years illustrate that it may be appropriate to place your work in conferences and journals in information science [50], in health science [51], and in mainstream health informatics [52].

**Discussion**

**Benefits of Developing and Using Conceptual Models**

This paper outlines current challenges in developing a conceptual model that integrates the information science aspects and the health science aspects of a health sector information and communication technology problem. The effort to overcome these challenges can yield important benefits that are not only theoretical but also practical. For the individual researcher, the conceptual model provides a persistent reminder of the defined entities and relationships that give shape and direction to their research plan and the interpretation of their findings. For diverse researchers with different questions who are formulating their own approach or discussing their own findings, the conceptual model facilitates comparisons and cross-pollination of ideas. Beyond the research community, there are benefits from giving closer attention to health informatics conceptual models for three professional practice communities: health informatics practitioners, other health practitioners, and other IT practitioners.

The community of health informatics professionals can use conceptual models more overtly to improve practice. By eliciting organizational input into conceptualizing implementation issues [13], they may be able to communicate and surmount notable problems of sustaining health information and communication technology applications [30]. For example, making deliberate use of a socio-technical conceptual model can help to anticipate unintended consequences before these emerge during system implementation [53]. Building health informatics conceptual modeling skills can enhance training and professional development for roles such as chief information officers in health organizations, research and development managers in health technology companies, and health informatics experts in large consulting firms. For example, health informaticians in such roles may benefit from working with conceptual models to explain and deliver the business value of information technology in healthcare [54].
Health professionals who are not information scientists can use conceptual models to facilitate inter-professional practice, exert collegial influence, and advance professional ethics in their work in an increasingly technological sector. For the individual health professional, making explicit the conceptual models you use in your professional practice enables you to integrate these more strongly into planning for new work practices during periods of technology change and adoption [55]. In working with colleagues in your profession, your ability to communicate conceptual models that frame health information and communication technology projects can position you as a leader and facilitator in the design and oversight of such projects [36]. In addition, the way each health profession expresses its ethical commitment to the safety and quality of care can be subtly different. When you clarify the conceptual models that you apply in health information technology projects that involve your own patients and/or clients in the settings where you provide care for them [56], you express more deeply the way you think about their needs and how technology interventions might address these. In doing so, you contribute to broadening the discourse within your profession about the ethical practice of clinical informatics.

Information professionals who are not health scientists can use integrative conceptual models to access important opportunities for innovation in the health sector, to achieve advances on health informatics grand challenges, and to build prized expertise and strong partnerships. Because information technology in health is developing massively and is being adopted rapidly, the health sector offers numerous opportunities to apply emerging solutions. When you aim to transfer into the health sector the conceptual models that underlie solutions from non-health sectors, acknowledging and articulating these enables you to reflect on how they can be expected to have positive impacts and be sustainable in health [57]. Coming to terms with the way information science conceptual models relate to the key health domain conceptual models where you are applying solutions can make you more effective in solving higher order problems [58]. If you are able to relate to and communicate within the health sector on this conceptual level [59,60], you will have expertise that is critical for successful collaboration with highly educated and committed health professionals to bring about technological transformation.

**Conclusion**

Our aim is to provide a representative selection of examples to accompany our suggestions for researchers who are new to the health informatics discipline. It is important for health informatics researchers to elucidate their hard-won conceptual modeling experience, even while recognizing that no conceptual model can ever be static or definitive. Sharing and comparing these foundations of knowledge can support good practice in research training (and in the commissioning, management, and review of research), and thereby can contribute to the evolution of better-formed and more strongly grounded health informatics research. Making explicit the defined entities and relationships in a health informatics research project facilitates deep engagement with cross-cutting problems, offers a way for researchers to be more effective, and enables research to have greater impact.

---

**Acknowledgments**

Ms Cecily Gilbert, Health and Biomedical Informatics Centre, University of Melbourne, did the initial literature review for this paper. JMIR reviewers’ feedback helped us greatly to sharpen the focus of this paper (reviewer names are acknowledged at the end of this paper).

**Authors’ Contributions**

KG conceived of the idea for the paper. KG and PS equally contributed to the structure and content of the paper and analyzed and added to the literature identified in the initial review. KG prepared the manuscript for submission and prepared responses to reviewers. PS approved the final version for submission.

**Conflicts of Interest**

None declared.

**References**


8. Michel G, Marcy T, Shiffman RN. A wireless, handheld decision support system to promote smoking cessation in primary care. AMIA Annu Symp Proc 2005:530-534 [FREE Full text] [Medline: 16779096]


©Kathleen Gray, Paulina Sockolow. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 24.02.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.
Integration of Provider, Pharmacy, and Patient-Reported Data to Improve Medication Adherence for Type 2 Diabetes: A Controlled Before-After Pilot Study

Abstract

Background: Patients with diabetes often have poor adherence to using medications as prescribed. The reasons why, however, are not well understood. Furthermore, most health care delivery processes do not routinely assess medication adherence or the factors that contribute to poor adherence.

Objective: The objective of the study was to assess the feasibility of an integrated informatics approach to aggregating and displaying clinically relevant data with the potential to identify issues that may interfere with appropriate medication utilization and facilitate patient-provider communication during clinical encounters about strategies to improve medication use.

Methods: We developed a clinical dashboard within an electronic health record (EHR) system that uses data from three sources: the medical record, pharmacy claims, and a patient portal. Next, we implemented the dashboard into three community health centers. Health care providers (n=15) and patients with diabetes (n=96) were enrolled in a before-after pilot to test the system’s impact on medication adherence and clinical outcomes. To measure adherence, we calculated the proportion of days covered using pharmacy claims. Demographic, laboratory, and visit data from the EHR were analyzed using pairwise t tests. Perceived barriers to adherence were self-reported by patients. Providers were surveyed about their use and perceptions of the clinical dashboard.

Results: Adherence significantly and meaningfully improved (improvements ranged from 6%-20%) consistently across diabetes as well as cardiovascular drug classes. Clinical outcomes, including HbA1c, blood pressure, lipid control, and emergency department utilization remained unchanged. Only a quarter of patients (n=24) logged into the patient portal and completed psychosocial questionnaires about their barriers to taking medications.

Conclusions: Integrated approaches using advanced EHR, clinical decision support, and patient-controlled technologies show promise for improving appropriate medication use and supporting better management of chronic conditions. Future research and development is necessary to design, implement, and integrate the myriad of EHR and clinical decision support systems as well as patient-focused information systems into routine care and patient processes that together support health and well-being.

doi:10.2196/medinform.4739
Introduction

Type 2 Diabetes Mellitus

Type 2 diabetes mellitus (T2DM) is a major public health issue, affecting more than 350 million people worldwide and the fourth leading cause of death. Globally, the prevalence of T2DM continues to rise at nearly epidemic rates, driven by urbanization, growing increases in obesity, and aging of populations [1]. Findings from several studies investigating the quality of T2DM care reveal a discrepancy between system-level disease management strategies and outcomes [2-6]. In essence, even though there are improved treatment strategies, expected outcomes are not occurring at a commensurate level. Therefore, greater emphases on patient-level factors that may explain T2DM intervention outcomes are being explored.

Patient Adherence to Medication

An example of a patient-level factor is adherence to complex medication regimens. Increasing evidence suggests that patients with T2DM often have poor adherence with prescribed medication therapies [7,8]. The reasons why individual patients do not take their medications as prescribed, however, are poorly understood. Existing research about medication adherence tends to investigate the issue as a class phenomenon, suggesting that patients as a group are universally impacted by a somewhat narrow range of factors such as side effects, cost, and forgetfulness [9].

Because adherence is treated as a class phenomenon, interventions tend to focus on singular modalities to change provider or patient behavior. This is especially true for informatics-related interventions. For example, in Vollmer et al [10], an interactive voice response system called patients who appeared to have gaps in refilling their asthma medication. Automated calls were made to patients or family members, but no assessment of individual barriers to adherence were measured or factored into the system. Similarly, recent systematic reviews of consumer-focused health information technologies conclude that prior studies tend to offer patients narrowly scoped functionalities out-of-the-box without regard to individual situations with limited success [11,12]. A recent review by Sapkota et al [13] found that while 22 of 52 (42%) interventions resulted in modest improvements in adherence, just 9 (17%) improved both adherence and glycemic control.

Despite existing efforts to improve adherence, patient-reported barriers to medication adherence (eg, lack of ability to pay, beliefs about the efficacy of medications in treating a condition, transportation to pharmacy, etc) and the extent to which those barriers contribute to poor T2DM outcomes are not currently assessed routinely in clinical practice [14]. Indeed, few have assessed the role of barriers perceived by patients to medications use and how perceived barriers may be addressed by intervention.

To address barriers to taking medications as prescribed facing individuals with T2DM, we developed a Web-based module for an electronic health record (EHR) system to electronically integrate the capture and presentation of information regarding T2DM patients’ disease management, medication adherence, and perceived barriers to adherence [15]. The system combines three elements: (1) objective data regarding medication possession ratios; (2) laboratory and point-of-care testing data that can indicate medication use; and (3) patient-reported data on perceived barriers to adherence. By routinely capturing patient-reported barriers and integrating such information with other electronic health data that is accessible during the clinical encounter, we seek to better potentiate patient-provider communication about medication use and thus inform T2DM therapy decision-making processes.

Following the development of the EHR module, we pilot tested the intervention in 3 primary care clinics located in an urban environment. By pilot testing the module, we sought to evaluate: (1) the extent to which patients would be willing to provide data to inform provider-patient conversations about medication use; (2) the impact of the system on medication use as prescribed by providers; and (3) the system implementation in a real-world clinical setting. We hypothesized that patients would be willing to share their perceived barriers to adherence as increasingly health systems are seeking patient input into clinical decision-making processes. We further hypothesized, given known challenges with adherence among patients with diabetes, that the intervention would improve adherence rates, which would improve other clinical indicators such as diabetes and cardiovascular risk factor control. Finally, we hypothesized that the dashboard would be used and positively perceived by providers.

Study Aim

In this paper, we describe the results of the pilot testing. We further comment on the implications of the pilot for future research and development of integrated informatics solutions to support patient-centered care.

Methods

Setting

Eskenazi Health is one of the 5 largest safety net health systems in the United States. The health system contains a 315-bed hospital and 9 community health centers located across the metropolitan area of Indianapolis, the eleventh largest city in the United States. Each community health center provides adult primary care, pediatrics, obstetrics, gynecology, and mental health.

The lead author (BED) presented the design of the dashboard and study to health center leadership at a health system meeting, asking for volunteers. In this study, 5 of the 9 community health centers volunteered to participate. We purposely selected 3 of the 5 health centers to ensure geographical and socioeconomic diversity.

http://medinform.jmir.org/2016/1/e4/
diversity among the study patient population. Only 3 health centers were selected to best manage provider and patient enrollment.

**Study Participants**

First, we recruited primary care providers with the authority to prescribe medications, which included medical doctors and nurse practitioners practicing in the 3 target clinics. Staff at ResNet, a practice-based research network recruitment engine used by Eskenazi as well as the Indiana Clinical and Translational Sciences Institute [16], approached eligible providers (n=29). Once providers gave their permission to participate, ResNet used the EHR to identify potential patients with diabetes for recruitment.

Provider participation was voluntary. Out of 29 eligible providers, 15 (52%) agreed to participate. Of the 14 eligible providers who did not participate, 8 never responded to our invitations to participate, 2 indicated they did not treat patients with T2DM, 1 indicated she was leaving the practice within 30 days, and 1 indicated she was not interested.

Potential patient participants were identified based on past medical history documented in the EHR. Queries of the EHR identified 2369 potential patients who were at least 18 years of age and possessed: (1) either an International Classification of Diseases, Version 9 or local diagnosis term indicating type 2 diabetes; (2) at least one active prescription for either biguanides, sulfonylureas, or thiazolidinediones (diabetic drug therapies); and (3) a history of primary care visits with an enrolled provider. ResNet staff then contacted eligible patients via phone to further screen them for enrollment. Screening questions asked patients to confirm they did have diabetes and were taking medication to manage their disease. In addition, phone screeners asked whether the patient possessed regular access to a computer with the Internet, as well as their ability to read English. After verification of inclusion criteria, phone screeners next described the study, and then asked for informed consent. Consenting patients were asked to provide either an email address or mobile phone number that could receive a text message to receive further instructions during the study. ResNet staff worked to contact, screen, and consent eligible patients until a sufficient number (n=96) was enrolled.

**System Description**

The primary system of interest is a clinical dashboard (Figure 1 shows this) used by providers in the context of routine primary care. The dashboard is a Java-based module designed to plug into the Regenstrief CareWeb framework, an open-source EHR platform developed by the Regenstrief Institute’s Center for Biomedical Informatics. CareWeb is a Web-based version of the Regenstrief Medical Record System (RMRS), providing primary care clinicians in Eskenazi Health facilities access to patients’ medical records [17]. We have previously described the design and development of the dashboard [15].

When a clinician selects a patient in the EHR, the dashboard refreshes with content from multiple sources: recent physiological data from the EHR, pharmacy data providing objective medication adherence data from a medication module, and patient-reported barriers to medication adherence from a patient portal. Briefly, we review the dashboard components and data sources.

**Figure 1.** Screenshot of clinical dashboard designed to integrate medication adherence information from multiple sources into electronic health record (EHR) system.

**Electronic Health Record Data**

The left panel of the dashboard displays information extracted from the RMRS. There are three types of EHR data relevant to diabetic patient populations that are displayed: blood pressure, HbA1c, and cholesterol. Clinicians routinely measure and analyze HbA1c levels to determine how well diabetic patients keep their disease under control. High blood pressure and poor lipid management are also commonly correlated with type 2 diabetes.
diabetes, leading to polypharmacy regimens to treat multiple conditions.

**Medication Adherence Data**

Adherence to T2DM and related medications is displayed in the middle panel of the dashboard. The information originates from the Medication (Med) Hub, an independent Web service within the Regenstrief technology infrastructure designed to gather and reconcile a patient’s current medication list [18]. Using available pharmacy data from the Med Hub, we calculate the proportion of days covered (PDC), a ratio representing whether the patient possessed a drug or a class of drugs (eg, all oral T2DM medications) during a defined measurement period. The PDC has been shown in numerous studies to accurately identify patients who fail to fill or refill their medications as directed by their physician or pharmacist [19], and it is the recommended measure for adherence by the Pharmacy Quality Alliance [20]. We use a dichotomized 6-month (180-day) PDC with a cut-off point of 80%, which we have found to provide the strongest and most reliable correlation with patient glycemic control [21].

**Patient-Reported Barriers to Adherence**

The right panel of the dashboard displays patient-reported barriers to taking their medications as prescribed. Patients report their barriers to adherence using a Web-based portal (Figure 2 shows this welcome screen) developed using the Open Medical Record System (OpenMRS) platform [22], an open-source EHR that originated at Regenstrief, but is now implemented and supported by a worldwide collaborative involving individuals from numerous counties involved in EHR and m-Health initiatives [23]. OpenMRS includes a forms module that allows collection of standardized data from patients. Using the forms module, we implemented a 5-point Likert style, validated questionnaire developed by researchers at the Diabetes Translational Research Center affiliated with the Indiana University School of Medicine [24,25]. The questionnaire, as implemented in the portal, is included as Appendix 1 (see Multimedia Appendix 1).

The patient questionnaire uses 20 items to assess possible barriers to medication adherence. For example, valid responses as to why one may not take his or her prescribed medications include, “I can’t afford them” and “I just forget to take them”. There are 5 factors or subscales that can be identified from responses to the questionnaire and displayed to clinical users: poor access to medications; poor communication with providers; poor understanding of medications and difficulty taking them or difficulty in taking them; presence of side effects; and system-level barriers to use. Previous analysis suggests that persons with poor cardiovascular disease (CVD) risk factor control have more reported barriers that may inhibit medication adherence than do persons with good risk factor control [25].

**Information Flow**

The information needed for the dashboard is queried in parallel with other CareWeb processes when a clinician opens the EHR for a patient (Figure 3 shows this). First, the CareWeb server notifies the T2DM module that a patient record has been selected. The module then, in parallel, requests data from the RMRS, Med Hub, and Web-based patient portal. The respective datasets are stored in the server’s cache until the clinician selects the “DM2 Med Adherence” tab within the CareWeb apps (Figure 1). Upon selection, the datasets are rendered into their respective columns for review by the clinician.
Figure 3. Information flow diagram depicting the architecture of the type 2 diabetes clinical information system module and its integration with existing electronic health record system, clinical decision support system, and patient portal components. T2DM = type 2 diabetes mellitus; EHR = electronic health record; CDS = clinical decision support; and Med = medication.

Study Design

Following enrollment of providers and patients, a research assistant created patient portal user accounts for each of the enrolled patients. Account information was then emailed or texted to patients, and they were asked to complete a baseline questionnaire regarding perceived barriers to using medications as prescribed by their clinician. Patients who did not complete a baseline questionnaire were reminded via email or text message, based on patient preferences. Patients were further asked to complete a questionnaire every 2-3 months after baseline. Reminders were sent to unresponsive participants. After successful completion of each questionnaire, patients were provided a US $10 retail gift card.

Consenting providers were given access to the integrated clinical dashboard within their EHR. This meant the tab labelled “DM2 Med Adherence” in Figure 1 was enabled for their user profile in the EHR, allowing them to access it at any time. Providers were informed about the tab and offered a brief tutorial by a nurse informaticist who regularly visits the clinic to educate providers about changes to the EHR. In addition, noninterruptive reminders were displayed to enrolled providers when relevant data for a patient were available for review on the dashboard. Providers could open the dashboard by clicking on the reminder in addition to, or instead of, clicking on the “DM2 Med Adherence” tab shown in Figure 1.

To assess the effect and implementation of the module, we used a controlled before-after design. We evaluated: (1) the willingness of patients to provide data via the Web-based portal; (2) changes to patients’ medication adherence as well as clinical indicators; and (3) provider usage and perceptions of the module. To measure patient willingness to provide data, we captured data on patient enrollment in the study as well as completion rates of adherence barrier questionnaires during the pilot timeframe. To measure changes in medication adherence as well as clinical indicators, we compared patients’ adherence, diabetic control, lipid control, and health care utilization rates before and after the introduction of the dashboard. Baseline data were collected from participants’ medical records for the year prior to the introduction of the intervention. The same data were collected from participants’ medical records for the 9-month pilot study. To measure provider engagement with the EHR system, we captured data on whether and how often they accessed the dashboard. We further surveyed providers about their use and perceptions of the dashboard following the pilot. The study was reviewed and approved by the Institutional Review Board at Indiana University (Protocol No. 1109006851).
Data Analysis

Descriptive statistics were calculated for participant characteristics as well as adherence, diabetic control, lipid control, and health care utilization variables. Mean values were calculated across all observations (e.g., multiple HbA1c measurements) during the preintervention (one year before) and postintervention (9 months after) periods, and the means for each time period were compared using within-subject paired t-tests. Adherence was measured using PDC calculated for each time period by patient for each drug class using the methods described in Nau [20] and Wang et al [26]. Patient-level PDC calculations were compared using within-subjects paired t-tests. Patient-reported barrier data from questionnaires were summarized using descriptive statistics. Provider responses to questionnaires regarding their use and perception of the dashboard were summarized using counts and means; small numbers prevented the use of statistical analysis. All statistical tests were performed using SAS 9.4 (Carey, NC).

Results

Study Recruitment

Out of 2369 potential patients identified by the EHR system, we attempted to recruit 906 (38.24%) via telephone in an effort to reach our goal of 100 enrolled patients. A total of 203/906 (22.40%) patients completed screening, of which 131/203 (64.5%) were eligible. Those determined to be ineligible (n=72/131; 55.0%) reported that they did not have regular access to a computer or mobile device with access to the Internet or could not provide informed consent. All of the potentially eligible participants were patients with diabetes who were taking medications as forecasted by the EHR. Of the 131 patients eligible to participate, 108 (82.4%) consented to participate in the study and 96 (73.3%) completed enrollment procedures.

Study Population Characteristics and Baseline Measures

Table 1 summarizes the characteristics of the final study population. African Americans were overrepresented given the population demographics of the Indianapolis metropolitan area from which they were selected. Most participants (n=84/96; 87%) were under 65, and half (n=50/96; 52%) possessed a baseline HbA1c above 8.0%, indicating they had difficulty controlling their diabetes. Participants were, on average, obese and possessed optimal low-density lipoprotein (LDL) cholesterol levels. On average, the participants visited their PCP (primary care provider) once every 3 months (mean >5 visits), and participants visited the emergency department (ED) once in the time period prior to the start of the intervention.

Table 1. Study population characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Count (%) or mean (median)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (%)</td>
<td>40 (42)</td>
<td></td>
</tr>
<tr>
<td>Female (%)</td>
<td>56 (58)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian (%)</td>
<td>47 (49)</td>
<td></td>
</tr>
<tr>
<td>African American (%)</td>
<td>41 (43)</td>
<td></td>
</tr>
<tr>
<td>Unknown (%)</td>
<td>8 (8)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>53 (52)</td>
<td>11.00</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>8.79 (8)</td>
<td>1.98</td>
</tr>
<tr>
<td>LDL (mg/dL)</td>
<td>95.6 (93)</td>
<td>34.18</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>39.87 (37)</td>
<td>11.95</td>
</tr>
<tr>
<td>PCP visits</td>
<td>5.45 (5)</td>
<td>4.71</td>
</tr>
<tr>
<td>ED visits</td>
<td>1.02 (1)</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Postintervention Change to Adherence and Other Measures

Table 2 summarizes the change in physiologic, health care utilization, and PDC calculations observed 9-months after the introduction of the intervention. Participants were repeatedly invited to log in to the Web-based portal to complete the questionnaire about challenges they faced in taking their medications. Despite multiple prompts via email and short message service, only 24 participants completed at least one questionnaire during the study period. We therefore stratified the results based on whether the patient completed the questionnaire. However, we observed no significant differences in demographics (e.g., gender, race) between patients who completed questionnaires and those who did not. We further observed little meaningful differences in clinical outcomes or health care utilization between those who did log in to the portal versus those who did not. However, medication adherence rates improved significantly and meaningfully across diabetes and cardiovascular drug classes.
Table 2. Patient engagement and outcomes measures.

<table>
<thead>
<tr>
<th>Outcomes, means</th>
<th>Assessment completed&lt;sup&gt;a&lt;/sup&gt; (n=24)</th>
<th>No assessment completed&lt;sup&gt;b&lt;/sup&gt; (n=24)</th>
<th>Overall total (n=96)</th>
<th>Overall P&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physiological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA1c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>8.71</td>
<td>8.82</td>
<td>8.79</td>
<td>.29</td>
</tr>
<tr>
<td>Post</td>
<td>8.66</td>
<td>8.96</td>
<td>8.88</td>
<td></td>
</tr>
<tr>
<td>LDL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>100.55</td>
<td>93.95</td>
<td>95.61</td>
<td>.06</td>
</tr>
<tr>
<td>Post</td>
<td>100.71</td>
<td>109.29</td>
<td>107.27</td>
<td></td>
</tr>
<tr>
<td>Body mass index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>40.41</td>
<td>39.71</td>
<td>39.87</td>
<td>.29</td>
</tr>
<tr>
<td>Post</td>
<td>44.43</td>
<td>37.97</td>
<td>39.46</td>
<td></td>
</tr>
<tr>
<td><strong>Utilization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>5.46</td>
<td>5.44</td>
<td>5.45</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Post</td>
<td>1.75</td>
<td>2.42</td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>ED visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>1.00</td>
<td>1.03</td>
<td>1.02</td>
<td>.001</td>
</tr>
<tr>
<td>Post</td>
<td>0.46</td>
<td>0.64</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td><strong>Diabetes drug classes, PDC %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDC for biguanides (n=38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>72</td>
<td>74</td>
<td>74</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Post</td>
<td>88</td>
<td>89</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>PDC for thiazolidinediones (n=8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>88</td>
<td>69</td>
<td>77</td>
<td>.004</td>
</tr>
<tr>
<td>Post</td>
<td>96</td>
<td>91</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>PDC for sulfonylureas (n=26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>71</td>
<td>74</td>
<td>73</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Post</td>
<td>88</td>
<td>96</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td><strong>Cardiovascular drug classes, PDC %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDC for ACE inhibitors (n=45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>90</td>
<td>85</td>
<td>85</td>
<td>.04</td>
</tr>
<tr>
<td>Post</td>
<td>83</td>
<td>85</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td><strong>PDC for angiotensin II receptor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antagonists (ARB) (n=14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>87</td>
<td>78</td>
<td>80</td>
<td>.02</td>
</tr>
<tr>
<td>Post</td>
<td>96</td>
<td>92</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>PDC for calcium channel blockers (n=18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>92</td>
<td>85</td>
<td>87</td>
<td>.03</td>
</tr>
<tr>
<td>Post</td>
<td>96</td>
<td>94</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>PDC for beta blockers (n=30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

JMIR Med Inform 2016 | vol. 4 | iss. 1 | e4 | p.64
http://medinform.jmir.org/2016/1/e4/
Overall total (n=96)

<table>
<thead>
<tr>
<th></th>
<th>Assessment completeda (n=24)</th>
<th>No assessment completedb (n=24)</th>
<th>Overall total (n=96)</th>
<th>Overall Pb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>86</td>
<td>73</td>
<td>75</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Post</td>
<td>94</td>
<td>90</td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>

**PDC for 3-hydroxy-3-methyl-glutaryl CoA reductase Inhibitors (statins) (n=13)**

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Post</td>
<td>99</td>
<td>91</td>
</tr>
</tbody>
</table>

aPatients who completed at least one Web-based assessment
bPatients who never logged in to complete an assessment
cPaired t test P value, overall

**Barriers to Adherence**
The top 5 barriers, based on mean scores from the questionnaires completed by those patients who completed at least one questionnaire (n=24), are presented in **Table 3**. The mean scores for these items range from 5 (Rarely) to 6 (Sometimes). So while common, these items do not necessarily indicate the items impact adherence. The items range in nature from general to financial to possible side effects.

**Table 3. Commonly reported perceived barriers to adherence by patients.**

<table>
<thead>
<tr>
<th>Questionnaire item</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>I just don't like taking medicine in general.</td>
<td>5.69</td>
</tr>
<tr>
<td>I just forget to take them.</td>
<td>5.36</td>
</tr>
<tr>
<td>I can't afford them.</td>
<td>5.19</td>
</tr>
<tr>
<td>I ran out of medication before I could call or visit my doctor or nurse.</td>
<td>5.17</td>
</tr>
<tr>
<td>My medicines make me feel bad or have side effects I don't like.</td>
<td>5.15</td>
</tr>
</tbody>
</table>

**Feedback From Providers**
Out of 29 eligible providers practicing at the 3 clinics prior to the implementation of the dashboard, 15 (52%) volunteered to participate. At the end of the 9-month study, 12 (80%) of these providers still practiced at the 3 participating sites. Out of the 12 providers still practicing, 6 (50%) provided feedback via a postintervention questionnaire.

All of the responding providers reported being aware of the dashboard when asked about it one month following the end of the pilot period. However, only 4 of the 6 (66%) reported using the dashboard at least once during the pilot. The 2 providers who indicated they did not use the dashboard responded negatively in general to questions about clinical information systems. For example, both these clinicians disagreed with statements that technology is easy to use in their workplace. They were further less than agreeable that technology could support care coordination or inform clinical decisions.

When asked about whether the dashboard was useful for patient-provider conversations about adherence and helpful to improving medication adherence or helpful to improving medication adherence, half of the providers responded negatively (eg, Disagree) and half were neutral. The providers who responded negatively to questions about the dashboard’s usefulness provided interesting open-ended comments. A provider reported that she was “not confident” regarding the quality of the information on the dashboard. Another respondent indicated he wasn’t sure if the dashboard contained data from outside pharmacies. A third provider recommended that patients “always bring their pill bottles” for independent verification of what they are taking, regardless of what data are displayed on the dashboard.

**Discussion**

**Principal Results**
A dashboard designed to inform clinicians and stimulate provider-patient conversations about medication adherence for patients with T2DM was introduced into the EHR system at 3 busy ambulatory clinics. Clinicians were prompted to review the dashboard for patients enrolled in a feasibility pilot, then work with their patients to address medication adherence issues illuminated by the dashboard. Following the introduction of the intervention, enrolled patients’ medication adherence improved significantly and meaningfully achieving its primary outcome. However, no significant changes to clinical outcomes or health care utilization were observed. This may be the result of patients being exposed to the barriers questionnaire; an exposure that may have stimulated their consideration of the importance of using medications as prescribed. In this regard, it is noteworthy that few patients were willing to complete Web-based questionnaires about their challenges to taking medications as prescribed using a Web-based portal from their home computer or mobile device. This may reflect concerns about divulging to
their providers that they are not using medications as prescribed. The desire of patients to indicate positive medication use patterns even when in actuality they are not adhering to avoid social sanctioning has been noted in the medication adherence literature [5,8,13]. Provider usage was limited and perceptions were mixed. Therefore, while the intervention shows potential for addressing patterns of medication use, it will require further work, especially more effective implementation strategies, to more actively engage patients in contributing data and providers in using the data if broader impact beyond adherence is to be achieved.

Medication adherence improved following the intervention. For every drug class in which mean adherence was below 80% (the cut-off point considered “good” by the Physicians Quality Alliance) prior to the intervention, mean PDC was observed to be above 80% at the end of the study. Moreover, this trend was consistent for diabetes as well as CVD drug classes, for which adherence was also displayed on the dashboard given known comorbidity between diabetes and CVD. In fact, PDC improved not only in the overall cohort, but also across subgroups with one exception: PDC fell for patients taking ACE inhibitors who completed at least one Web-based questionnaire, although it did not dip below 80%. Thus, the dashboard appears to be effective at calling attention to adherence issues, even in cases where the dashboard lacked information on patients’ self-reported psychosocial challenges to taking their medications.

The combination of providing PDC alongside laboratory data may therefore be sufficient to stimulate patient-provider discussion about appropriate medication use. It is also possible that the requests for patients to provide data about barriers acted as a clinical reminder, which influenced patients to be mindful about medication use. Finally, it is plausible that patients, believing that their provider was going to see data about their medication utilization, elected to be more consistent in their medication use. Information on patient-provider conversations and motivations for better adherence were not captured in this study; these would be important dimensions to measure in a larger study involving mixed methods.

Although the primary objective was achieved, we did not observe any meaningful changes to patients’ health status. Diabetic control and BMI remained unchanged, and health care utilization remained constant. With respect to diabetic control, in theory, this should improve in parallel with medication adherence. Where patients with diabetes use their medications as directed by clinicians, HbA1c levels are expected to fall below 8.0%. A reason we may not have observed a change is that the requests for patients to provide data about barriers acted as a clinical reminder, which influenced patients to be mindful about medication use. Finally, it is plausible that patients, believing that their provider was going to see data about their medication utilization, elected to be more consistent in their medication use.

In addition, the project struggled to engage patients in logging into the portal where they could complete the psychosocial questionnaire about barriers to taking their medications. Even with a financial incentive, just one-quarter of the cohort completed at least one questionnaire during the 9-month pilot. Many enrolled patients struggled due to computer or Internet access issues. For example, one participant repeatedly stated she was waiting for her daughter to come over to help her. Other patients’ email addresses bounced less than 1-2 weeks after they provided them during the enrollment process. Text messaging and phone calls were helpful in reaching some of these patients, but access challenges remained for a significant number of study subjects. Furthermore, the patient portal was not routinely used by our clinical partner, Eskenazi Health. The portal vendor for the health system refused to work with the study team, stating their platform was designed only for secure messaging and they did not have an interest in expanding their service offerings to enable patient reported data to be integrated into the EHR.

In addition, the desire of patients to indicate positive medication use may have contributed to lackluster participation by patients beyond their Internet access challenges.

Comparison With Prior Work

Prior studies to improve adherence and glycemic control or glycemic control tend to focus on singular modalities to change provider or patient behavior with limited success. For example, in Vollmer et al [10], an interactive voice response system called patients who appeared to have gaps in refilling their asthma medication. The system was statistically significant in changing adherence, but the mean change (2%) was not clinically meaningful (eg, impact on health outcomes as well as quality of life). Similarly, a systematic review of patient portals by Ammenwerth et al [12] identified just one study that demonstrated an effect on diabetes care delivery. While that one portal was found to be associated with a change in medication regimen, it had no impact on clinical outcomes as measured by HbA1c and blood pressure [27]. A recent
systematic review by Sapkota et al [13] of interventions targeting patients with type 2 diabetes found that just 9 of 52 (17%) studies found an improvement to both adherence and glycemic control. A broader meta-analysis of consumer-focused health information technology trials by Or and Tao [11] similarly found that none of the identified studies showed an impact on clinical or psychosocial outcomes; and the impact of patient-centered technologies had mixed effects on consumer behavior. While our intervention did have both statistically significant and clinically meaningful impact on adherence, our impact on more core measures of quality were similarly disappointing.

**Future Directions**

Given the context of prior studies, our results suggest a change to how we design, implement, and integrate technologies into care delivery systems and consumers’ lives. Our dashboard received positive feedback from some of the clinicians before and during the study. Yet to access the dashboard, clinicians needed to click on yet another tab within the EHR system or be prompted to view the dashboard when data were available for a given patient. While alerts can be useful to guide provider behavior, they can also be viewed as a nuisance leading to providers ignoring or overriding them [28,29]. This may be why some providers reported not using the dashboard during the pilot. Therefore, as we develop dashboards or other EHR widgets, we must find ways to more seamlessly integrate them into clinical processes when contextually appropriate. Other projects at Regenstrief have explored contextually sensitive alerts [30-32], concluding they are promising. Furthermore, we need to explore other clinical team roles (eg, medical assistant, registered nurse, clinical pharmacist), beyond the physician, for whom an adherence dashboard might make more sense given the workflow in primary care as well as other settings.

In addition to information and workflows, we can refine our approach to synthesis as well as visualization of information in clinical dashboards. For example, in this study, patient reported barriers were often missing, yet medication adherence improved. Therefore, it may be sufficient to provide a view of adherence using just 2, objective data sources (pharmacy claims, vital signs). This would assume, however, that providers trust the data and dashboard. In this feasibility pilot, 2 of the providers were wary of the data in the dashboard. Better awareness and experience with integrated data views might solve this issue; or better strategies to educate providers on the sources and quality of data in EHR systems might be considered.

Moreover, when patient barriers were missing, a full third of the screen was blank. Other visualizations of adherence data in combination with vitals sign trends and patient-reported data or patient-reported data could be explored to design EHR widgets that might reduce the potential for white space or maximize screen real estate. A larger trial could explore, for example, variants to the information visualization tested here to find optimal representations of multiple data streams.

To be effective, “routine clinical care” must also include environments where patients exist (eg, home, work, bus stop). Integrated information solutions, therefore, must incorporate technologies that can be integrated into daily routines of people. Our efforts were hampered by system barriers that remain a challenge for many people, especially those who are elderly or of low socioeconomic status [33]. Thus, a digital divide still exists, even if it may be shrinking in populations burdened by multiple comorbid diseases requiring complex drug regimens. Therefore, future studies should consider approaches that can reach broad populations instead of developing on an isolated platform that cannot be integrated into the existing health information infrastructure. A pilot study involving text messages to homeless veterans that reduced appointment no-show rates [34] and the multiple studies involving text4baby [35] demonstrate that even disadvantaged populations can benefit from health informatics interventions delivered in a modality that is broadly accessible to patients. Instead of the Web-based browser-based portal we used, we could have explored mobile platforms that could be accessed via mobile phones. Yet one-dimensional “smart” apps are similarly unlikely to be sufficient. The patient-focused aspect of our pilot was largely one-sided, designed to collect information from patients instead of engage them in learning about their disease and strategies for self-management. Management of diabetes involves not only medication adherence, but also changes to diet, exercise, and comorbid conditions such as hypertension. Therefore, technologies that engage patients will likely be those that can address multiple concerns in an integrated fashion (eg, one-stop shop) rather than require multiple apps or interfaces. In addition, it will become critical that future apps include appropriate education that emphasizes the importance of patient-provider communication to enable shared decision making that will achieve optimal therapeutic outcomes.

**Conclusions**

To fully realize the potential of health information technologies to support patient-centered care delivery, while impacting population health outcomes, information and technical systems need to be integrated. Our vision of technical interfaces among EHR, CDS, and patient information systems, which are integrated into clinical and personal ecosystems, is necessary to create environments in which shared decision making can be informed by evidence, an individual’s health data, and knowledge of social determinants. The results from our early pilot of an integrated approach are promising. Yet, they suggest additional research and development to better design, implement, and integrate the myriad of EHR and CDS systems, as well as patient devices, into routine care and patient processes that together support health and well-being.

**Acknowledgments**

The authors thank Jeremy Leventhal, Jonathan Cummins, and Hui Xiao of the Regenstrief Institute Center for Biomedical Informatics for their time and talent in architecting and developing the systems described in this article. Zuoyi Zhang, PhD, of the Regenstrief Institute provided invaluable advice and guidance on the analysis plan. The authors also wish to acknowledge
Abdulrahman M. Jabour, a doctoral candidate at the IU School of Informatics and Computing, for his assistance in managing study subject’s portal accounts and collection of data during the early phases of the pilot study. The authors further acknowledge the support of ResNet, especially Jane Anne French, CCRP, in recruiting patients from Eskenazi Health for enrollment in this study. ResNet is a core service of the Indiana Clinical and Translational Sciences Institute that provides recruitment services. Finally, the authors thank the clinical leadership at Eskenazi Health for their support in finding innovative ways to integrate informatics tools into clinical practice.

This work was supported by a grant (R34DK092769) from the National Institute for Diabetes and Digestive and Kidney Diseases. The Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (CIN 13-416) further employs Dr Dixon as a Health Research Scientist at the Richard L. Roudebush Veterans Affairs Medical Center in Indianapolis, Indiana, USA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Diabetes and Digestive and Kidney Diseases, the National Institutes of Health, the Department of Veterans Affairs, or the U.S. Government.

Authors' Contributions

DGM and BED guided the design of the study. AHA provided key analytical support for calculating and summarizing the medication adherence data. EOP provided key support for the recruitment and enrollment of patients as well as data management and analysis. BED drafted the article. DGM, AHA, and EOP provided critical revisions of the article. BED finalized the article for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire given to patients regarding perceived barriers to taking their medications as prescribed.

References


25. Marrero D, Monahan P, Lane K, Hayes R. Validation of a scale to measure patient-perceived barriers to medication use. 2006 Presented at: International Society for Quality of Life Research meeting abstracts. Qual Life Res. 15(Suppl 1), Abstract #1223; 2006; USA p. A34-A35 URL: https://scholarworks.iupui.edu/handle/1805/8164


Abbreviations

BMI: body mass index
CVD: cardiovascular disease
ED: emergency department
EHR: electronic health record
LDL: low-density lipoprotein
MED: medication
OpenMRS: Open Medical Record System
PCP: primary care provider
PDC: proportion of days covered
RMRS: Regenstrief Medical Record System
T2DM: type 2 diabetes mellitus

©Brian E Dixon, Abdullah H Alzeer, Erin O’Kelly Phillips, David G Marrero. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 08.02.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.
Industry and Occupation in the Electronic Health Record: An Investigation of the National Institute for Occupational Safety and Health Industry and Occupation Computerized Coding System

Matthew Schmitz\*, MSc; Linda Forst\*, MD, MPH
Environmental and Occupational Health Sciences, School of Public Health, University of Illinois at Chicago, Chicago, IL, United States
*all authors contributed equally

Corresponding Author:
Linda Forst, MD, MPH
Environmental and Occupational Health Sciences
School of Public Health
University of Illinois at Chicago
2121 W. Taylor, MC 922
Chicago, IL, 60612
United States
Phone: 1 312 3553826
Fax: 1 312 413 9898
Email: Forst-L@uic.edu

Abstract

Background: Inclusion of information about a patient’s work, industry, and occupation, in the electronic health record (EHR) could facilitate occupational health surveillance, better health outcomes, prevention activities, and identification of workers’ compensation cases. The US National Institute for Occupational Safety and Health (NIOSH) has developed an autocoding system for “industry” and “occupation” based on 1990 Bureau of Census codes; its effectiveness requires evaluation in conjunction with promoting the mandatory addition of these variables to the EHR.

Objective: The objective of the study was to evaluate the intercoder reliability of NIOSH’s Industry and Occupation Computerized Coding System (NIOCCS) when applied to data collected in a community survey conducted under the Affordable Care Act; to determine the proportion of records that are autocoded using NIOCCS.

Methods: Standard Occupational Classification (SOC) codes are used by several federal agencies in databases that capture demographic, employment, and health information to harmonize variables related to work activities among these data sources. There are 359 industry and occupation responses that were hand coded by 2 investigators, who came to a consensus on every code. The same variables were autocoded using NIOCCS at the high and moderate criteria level.

Results: Kappa was .84 for agreement between hand coders and between the hand coder consensus code versus NIOCCS high confidence level codes for the first 2 digits of the SOC code. For 4 digits, NIOCCS coding versus investigator coding ranged from kappa=.56 to .70. In this study, NIOCCS was able to achieve production rates (ie, to autocode) 31%-36% of entered variables at the “high confidence” level and 49%-58% at the “medium confidence” level. Autocoding (production) rates are somewhat lower than those reported by NIOSH. Agreement between manually coded and autocoded data are “substantial” at the 2-digit level, but only “fair” to “good” at the 4-digit level.

Conclusions: This work serves as a baseline for performance of NIOCCS by investigators in the field. Further field testing will clarify NIOCCS effectiveness in terms of ability to assign codes and coding accuracy and will clarify its value as inclusion of these occupational variables in the EHR is promoted.

(JMIR Med Inform 2016;4(1):e5) doi:10.2196/medinform.4839

KEYWORDS
medical informatics; occupation code; industry code; NIOCCS; occupational health; occupation+electronic health record
Introduction

The links between work and health have been long understood [1]. Illnesses and injuries that arise out of employment are amenable to primary, secondary, and tertiary prevention, but prevention requires identification of hazards as well as knowledge of associated health conditions. Inclusion of information about employment in medical records could aid physicians and other health professionals in the recognition of work-related illnesses and injuries. Such information could also highlight working conditions that interfere with general health and well-being and stymie treatment efforts. Furthermore, capture of employment variables would allow for intervention research and surveillance of work-related patterns of illness and injury at the population level [2-4]. After earlier attempts at automated coding of work variables for governmental agencies in the United States [5], the development and widespread use of electronic health records (EHRs) presents a unique opportunity to advance the collection and incorporation of information on industry and occupation (I&O) into patients’ medical records [6].

In its 2010 report, “Reducing Environmental Cancer Risk - What We Can Do Now”, the President’s Cancer Panel recommended routine assessment of occupational history and the incorporation of such information into the medical record in order to better assess potential workplace exposures and related risk of chronic illnesses such as cancer [7]. The National Prevention Council, headed by the US Surgeon General, also called for the inclusion of occupational and environmental risk assessment in the patient medical history in its 2011 National Prevention Strategy [8]. An Institute of Medicine 2011 report, “Incorporating Occupational Information in Electronic Health Records”, concluded that incorporation of occupational information in EHRs could contribute to improving individual and population health care and issued a number of recommendations to the US National Institute for Occupational Safety and Health (NIOSH) and its partners to guide the process; these include adopting Standard Occupational Classification (SOC) coding standards for use in EHRs, assessing the feasibility of autocoding occupational health information, and assessing the impact on meaningful-use goals of incorporating occupational information into EHRs, specific objectives that must be met to qualify for Centers for Medicare & Medicaid Services Incentive Programs meant to ensure that implementation of EHRs improves clinical outcomes and population health [9]. As a result of these recommendations, NIOSH developed the National Industry and Occupation Computerized Coding System (NIOCCS), and following on to earlier work in this area, a Web-based system that translates “industry” and “occupation” text into standardized codes. The NIOSCCS system has had limited evaluation, to date.

The goal of this investigation was to evaluate the quality of I&O coding of data obtained in a community-based, health care needs assessment. Specific objectives were to: (1) determine intrarater reliability of hand coding of “industry” and “occupation” variables using the 2010 Standard Occupational Codes; (2) determine the ability of NIOCCS to assign codes to “industry” and “occupation” responses from the general public (potential patients); and (3) to evaluate the relationship between hand-coded versus NIOCCS-assigned SOC codes.

Methods

University of Illinois Survey on Neighborhood Health

In 2013-2014, the University of Illinois Survey on Neighborhood Health (UNISON) was conducted, as this health care system’s community health needs assessment required by the Patient Accountability and Affordable Care Act (ACA) [10]. A sampling scheme was constructed for a study of 1400 individuals, approximately half of them reached through a door-to-door survey according to a randomized, block design; the other half that were surveyed were current patients with specific chronic diseases.

A survey tool was created from existing national surveys to collect patient-reported information about health behaviors, health care access and utilization, prevalence of disease conditions, quality of life indicators, and knowledge of the ACA. Basic biometric screening was done and those who answered the survey were invited to come to university clinics for laboratory testing. There were 3 questions that were inserted by investigators to use for the current study: (1) current employment status; (2) type of business or industry working in (respondents could select from a list or write in a response); and (3) job title (write-in, only). For every currently employed individual, responses to questions 2 and 3 were downloaded to a MS Excel spreadsheet and used in this analysis. No identifying information was obtained.

Coding

The 2010 SOC system is used to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data [11]. This system has been adopted by several federal agencies in order to harmonize collection of work-related variables in a variety of demographic, employment, and health oriented databases. The 2010 version classifies workers into one of 840 detailed occupations according to their definitions. To facilitate classification, detailed occupations are combined to form 461 broad occupations, 97 minor groups, and 23 major groups. Detailed occupations in the SOC with similar job duties, and in some cases skills, education, and training, are grouped together. The coding scheme gives 2 digits for the 23 major categories, followed by a hyphen, and 4 more digits for more finely described categories (XX-XXXX). For example, a home health aide would be coded Healthcare Support Occupations (31-0000) for major group, (31-1000) for minor group, Nursing, Psychiatric, and Home Health Aides (31-1010) for broad occupation, and Home Health Aide (31-1011) for detailed occupation [11].

The investigators obtained the 2010 SOC system of codes and individually hand-coded the data down to the most detailed level of classification possible based on the information provided by each respondent [11]. They then met and came to a consensus if their codes differed.

The NIOSH Industry & Occupation Computerized Coding System (NIOCCS) Version 2.0 was used to autocode the data according to 2010 SOC codes, separately from the hand coding.
NIOCCS is a Web-based system that translates text into standardized I&O codes [12]. It was designed for use by researchers, government agencies, state health departments, and other organizations that collect or evaluate information using I&O. Its purpose is to provide a tool that reduces the high cost of manually coding I&O information, while simultaneously improving uniformity of the codes. The NIOCCS Coding Engine design has processes that cover phrase-based and word-based, exact match and proximity match, and weighted and not-weighted matching. Each process has its specialty of best-fit coding areas to enhance the combined coding ability. Figure 1 shows the NIOCCS coding scheme [12].

The dataset of survey participants’ responses on employment was used for NIOCCS autocoding for the first 2 digits (industry) and the last 4 digits (occupation) using both the high confidence level (90% cutoff) and medium confidence level (70% cutoff) criteria. Records processed using the high confidence threshold require that NIOCCS has 90% or greater confidence of accuracy for matching, whereas the medium confidence threshold only requires 70% or greater confidence of accuracy for coding to occur. The degree of confidence is based on the degree of fidelity of the variable to the actual code, with the “high” confidence level requiring stronger probability of a match (less “fuzziness”) than “medium” confidence level allows. We used either the industry data (first 2 digits) selected by the respondent from a list, or the write-in response when “other” was selected. The occupation data, the last 4 digits, were written in by the respondent (not selected off a list).

Figure 1. The US National Institute for Occupational Safety and Health (NIOSH) Industry & Occupation Computerized Coding System (NIOCCS) Coding Engine [12]. I&O: industry and occupation.

Data Analysis

The kappa statistic was used to assess the reliability of manual versus NIOCCS autocoding. The first 2 digits were compared between hand and electronic coding, as were the last 4 digits. The proportion of records containing I&O information that are autocoded out of all the records submitted for coding is called the “production rate”. The industries and occupations that could not be coded by NIOCCS resulted in unbalanced contingency tables, a row by column table where the manually assigned codes for each respondent constitute the rows and the NIOCCS-assigned codes constitute the columns, for both the 2-digit and 4-digit comparisons. A contingency table is one where rows and columns are set up to analyze for associations between the two. In order to obtain correct kappa statistics, an approach suggested by Crewson [13] was used to create balanced contingency tables through the creation of dummy observations covering all possible rating scenarios in a separate stratum from the original data. Using this approach, SAS PROC FREQ calculates the kappa statistics separately for the strata containing the original data and the dummy observations. All 6 digits were not evaluated because of the extremely high number of potential rating categories and practical limits of computing power. SAS v.9.4 (SAS Inc, North Carolina) was used for this analysis.

Results

Response Breakdown From Door-to-Door Survey

Responses were obtained from 740 individuals in the door-to-door survey, with 359 (48.5%) currently working. The breakdown of “industry” based on hand coding, considered the
gold standard for I&O coding, is shown in Table 1. There were 6 of the 23 industries that comprised over 50% of the responses.

Table 1. Industrial sectors of 359 working individuals surveyed in the UNISON study.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Number</th>
<th>%</th>
<th>Cum %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office and administrative support (43-XXXX)</td>
<td>37</td>
<td>10.3</td>
<td>10.3</td>
</tr>
<tr>
<td>Management (11-XXXX)</td>
<td>32</td>
<td>8.9</td>
<td>19.2</td>
</tr>
<tr>
<td>Health care support (31-XXXX)</td>
<td>30</td>
<td>8.4</td>
<td>27.6</td>
</tr>
<tr>
<td>Food preparation and serving (35-XXXX)</td>
<td>29</td>
<td>8.1</td>
<td>35.7</td>
</tr>
<tr>
<td>Education, training, and library (25-XXXX)</td>
<td>28</td>
<td>7.8</td>
<td>43.5</td>
</tr>
<tr>
<td>Personal care and service (39-XXXX)</td>
<td>24</td>
<td>6.7</td>
<td>50.2</td>
</tr>
<tr>
<td>Transportation (53-XXXX)</td>
<td>23</td>
<td>6.4</td>
<td>56.6</td>
</tr>
<tr>
<td>Sales and related (41-XXXX)</td>
<td>20</td>
<td>5.6</td>
<td>62.2</td>
</tr>
<tr>
<td>Business and financial operations (13-XXXX)</td>
<td>18</td>
<td>5.0</td>
<td>67.2</td>
</tr>
<tr>
<td>Production (51-XXXX)</td>
<td>18</td>
<td>5.0</td>
<td>72.2</td>
</tr>
<tr>
<td>Protective service (33-XXXX)</td>
<td>14</td>
<td>3.9</td>
<td>76.1</td>
</tr>
<tr>
<td>Health care practitioners and technical (29-XXXX)</td>
<td>13</td>
<td>3.6</td>
<td>79.7</td>
</tr>
<tr>
<td>Arts, design, entertainment, sports, and media (27-XXXX)</td>
<td>11</td>
<td>3.1</td>
<td>82.8</td>
</tr>
<tr>
<td>Building and grounds cleaning and maintenance (37-XXXX)</td>
<td>11</td>
<td>3.1</td>
<td>85.9</td>
</tr>
<tr>
<td>Life, physical, and social science (19-XXXX)</td>
<td>9</td>
<td>2.5</td>
<td>88.4</td>
</tr>
<tr>
<td>Computer and mathematical (15-XXXX)</td>
<td>8</td>
<td>2.2</td>
<td>90.6</td>
</tr>
<tr>
<td>Community and social service (21-XXXX)</td>
<td>8</td>
<td>2.2</td>
<td>92.8</td>
</tr>
<tr>
<td>Construction and extraction (47-XXXX)</td>
<td>7</td>
<td>1.9</td>
<td>94.8</td>
</tr>
<tr>
<td>Job title “None”, left blank, or unknown</td>
<td>6</td>
<td>1.7</td>
<td>96.5</td>
</tr>
<tr>
<td>Legal (23-XXXX)</td>
<td>5</td>
<td>1.4</td>
<td>97.9</td>
</tr>
<tr>
<td>Installation, maintenance, and repair (49-XXXX)</td>
<td>4</td>
<td>1.1</td>
<td>99.0</td>
</tr>
<tr>
<td>Architecture and engineering (17-XXXX)</td>
<td>2</td>
<td>0.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Farming, fishing, and forestry (45-XXXX)</td>
<td>1</td>
<td>0.3</td>
<td>99.9</td>
</tr>
<tr>
<td>Temp worker</td>
<td>1</td>
<td>0.3</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Production Rates for “High” and “Medium” Confidence Levels

Table 2 shows the production rates (the percentage autocoded) for “high” and “medium” confidence levels. For the high confidence algorithm, the production rates of “industry” (the first 2 digits) range from 115/359 (32.0%) to 129/359 (35.9%), with the “write-ins” performing a little better (23/65, 35% autocoded) than those selected off a card (94/294, 31.9% autocoded). Production rates were higher for the medium confidence level (176/294, 59.8% to 38/65, 58%). Similar results were obtained for the production rates of “occupation” (the last 4 digits) at both confidence levels.

Table 2. Production rates of autocoding of I&O by 2- and 4- digit Standardized Occupational Codes obtained from 359 respondents in a community survey, with industry selected from a list versus industry write-in.

<table>
<thead>
<tr>
<th>SOC</th>
<th>High confidence, %</th>
<th>Medium confidence, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-digit (XX-XXXX), industry selected off card</td>
<td>32</td>
<td>49</td>
</tr>
<tr>
<td>2-digit (XX-XXXX), industry write-in</td>
<td>36</td>
<td>58</td>
</tr>
<tr>
<td>4-digit (XX-XXXX), industry selected off card</td>
<td>31</td>
<td>49</td>
</tr>
<tr>
<td>4-digit (XX-XXXX), industry write-in</td>
<td>36</td>
<td>58</td>
</tr>
</tbody>
</table>

Comparison of Manual Coding by Two Investigators

Both investigators manually coded all 359 responses. For the comparison of manual coding by each of the investigators, 2-digit blinded coding yielded a kappa of 0.84 (95% CI 0.79-0.88) using 338 observations (21 missing); at the 4-digit level, a kappa of 0.58 (95% CI 0.52-0.63) using 337 observations (22 missing) was achieved. Investigators subsequently reached
agreement at the 6-digit level (2 and 4, combined) on every case and used these agreed-upon codes for comparisons to the NIOCCS-generated codes. Hand coding versus NIOCCS coding yielded high kappa statistics at the 2-digit level, but low kappa statistics at the 4-digit level. The NIOCCS high confidence algorithm produced higher accuracy than the medium confidence level, but coded fewer observations. There was no difference in selection of industry off a card versus writing it in free form. All comparisons are shown in Table 3.

Table 3. Interrater reliability measures of SOC 2010 coding of I&O data by different coding methods.

<table>
<thead>
<tr>
<th>Coding technique</th>
<th>n (missing)</th>
<th>Kappa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 1 x Investigator 2: 2 digit</td>
<td>338 (21)</td>
<td>0.84a (0.79-0.88)</td>
</tr>
<tr>
<td>Investigator 1 x Investigator 2: 4 digit</td>
<td>337 (22)</td>
<td>0.58 (0.52-0.63)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS high: 2 digit</td>
<td>115 (244)</td>
<td>0.84a (0.77-0.91)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS high: 2 digit (with write-in)</td>
<td>129 (230)</td>
<td>0.84a (0.78-0.91)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS high: 4 digit</td>
<td>112 (247)</td>
<td>0.70 (0.62-0.79)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS high: 4 digit (with write-in)</td>
<td>129 (230)</td>
<td>0.68 (0.60-0.76)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS med: 2 digit</td>
<td>177 (182)</td>
<td>0.71 (0.65-0.78)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS med: 2 digit (with write-in)</td>
<td>207 (152)</td>
<td>0.71 (0.64-0.77)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS med: 4 digit</td>
<td>177 (182)</td>
<td>0.60 (0.52-0.67)</td>
</tr>
<tr>
<td>Investigator-agreed x NIOCCS med: 4 digit (with write-in)</td>
<td>207 (152)</td>
<td>0.56 (0.49-0.63)</td>
</tr>
</tbody>
</table>

a agreement = high

**Discussion**

**Relationship Between Work and Health**

There is a growing appreciation of the relationship between work and health. First, “work” is a determinant of health in that job activities expose working people to illness and injury risks: workers who are exposed to chemical, biological, physical, ergonomic, and psychosocial hazards are more likely to experience adverse health effects associated with those hazards. Furthermore, health is affected by employment status, whether an individual is employed, unemployed, partially employed, or stably employed has an impact on health and well-being [14]. Second, employment in the United States is integrally related to health insurance coverage for both general health and work-related injury/illness, and health care coverage determines whether an individual has access to care or whether optimal health outcomes can be achieved as a result of a clinical encounter. Third, the workplace is increasingly being utilized as a venue for health promotion activities. Wellness programs are being implemented to prevent and control chronic health conditions based on the belief that a healthy workforce is a more productive workforce and will cost the employer less money due to absenteeism, presenteeism (low productivity while at work), and health insurance costs [15-17]. Finally, there are recent studies examining the transfer of workers’ compensation costs to general health insurance, to federal programs, to community health centers as nonreimbursable costs, and to indigent worker-patients, themselves [18-20]. Inclusion of information on I&O in the EHR is integral to providing information that can be utilized in health care encounters to improve the health and well-being of adult patients. On a population level, variables collected in the EHR can be used to study relationships between sociodemographics, hazardous exposure conditions, and health outcomes. The use of health records by state programs for surveillance of occupational illness and injury has increased in the last decade as a way to enhance case capture, target preventive efforts, and conduct research [21].

General guidelines for the interpretation of the kappa statistic have been suggested by Landis and Koch [22] as values <.00 indicating poor agreement, .00-.20 as slight, .21-.40 as fair, .41-.60 as moderate, .61-.80 as substantial, and .81-1.00 as almost perfect agreement. Fleiss [23] has also suggested guidelines characterizing kappa values below .40 as poor, .40-.75 as fair to good, and over .75 as excellent. While both sets of guidelines are arbitrary, they provide a general framework for the interpretation of kappa values. It is also important to remember that the magnitude of kappa values can be influenced by the prevalence of given attributes being rated, in this case the prevalence of employment in various industries and occupations among the sample population, as well as potentially by bias [24].

**Principal Results**

Using the Landis and Koch [22] framework for interpreting the kappa statistic, the between-investigator agreement for 2-digit manual coding was “almost perfect”, as was the agreement between the manually coded classification and the NIOCCS high confidence model, while the agreement between the manually coded classification and the NIOCCS medium confidence model was slightly lower, but still “substantial”. The between-investigator agreement for 4-digit coding was moderate, the agreement between the manually coded classification and the NIOCCS high confidence model was substantial, and the agreement between the manually coded classification and the NIOCCS medium confidence model was moderate. Including respondent-provided write-in information for industry led to a slight increase in the number of observations.
autocoded by NIOCCS for both high and medium confidence models, but did not alter the kappa statistic for 2-digit coding and slightly decreased the kappa statistic for 4-digit coding as compared to the manually coded classifications.

If the investigator-agreed-upon manually coded classifications are taken as the true I&O classifications for this study, then the NIOCCS models performed near, but slightly below the accuracy performance goals of 10% or less error rate and 25% or less error rate for the high and medium confidence models respectively, when examining the 2-digit coding; however, it performed well below the performance goals for the 4-digit coding [25]. Importantly, the NIOCCS high confidence model achieved a comparable level of agreement for the 2-digit coding and a greater level of agreement for the 4-digit coding than was originally achieved between blinded investigators. However, a significant limitation of the model that may influence the overall agreement between manual and autocoded results is the production rate of the NIOCCS autocoding (ie, most of the entries do not get coded by NIOCCS and therefore cannot be compared with manual coding). Notably, there does not appear to be any pattern in regards to industry or occupation in the data that was not autocoded.

The goal of NIOCCS autocoding is to simplify and mechanize insertion of I&O variables into the EHR; production rate and accuracy are critical to streamlining inclusion of these variables. This study used a field-based survey to explore public use of NIOCCS for coding I&O. Data provided in Table 4 show that NIOCCS did not perform as well in this study as it has in other settings. Death certificates yield the highest production rates and coding by NIOSH performs better than it did for external investigators in this study.

### Table 4. NIOCCS autocoding production rates (proportion autocoded) for this investigation versus other studies.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Year 2013, %</th>
<th>Year 2014 (Jan-Sept), %</th>
<th>Illinois ACA survey, %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NIOCCS internal user (NIOSH personnel) results</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death certificates</td>
<td>60</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td><strong>NIOCCS external user (non-NIOSH personnel) results</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death certificates</td>
<td>64 [28]</td>
<td>64 [28]</td>
<td>32-36</td>
</tr>
<tr>
<td>Cancer registries</td>
<td>35 [28]</td>
<td>60 [28]</td>
<td></td>
</tr>
<tr>
<td>Surveys</td>
<td>49 [28]</td>
<td>50 [28]</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>52 [28]</td>
<td>57 [28]</td>
<td></td>
</tr>
<tr>
<td>Average—all data types</td>
<td>51</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>BRFSS-10 states [27]</td>
<td>31-55</td>
<td></td>
<td>(avg=42%)¹</td>
</tr>
</tbody>
</table>

¹ Personal communication, NIOSH

### Production Rates of the National Industry and Occupation Computerized Coding System

The production rates of the NIOCCS high confidence model were 32% and 36% for the selected and write-in augmented data, respectively; the production rates using the medium confidence model were 49% and 58% for the selected and write-in augmented data. The observed production rates are below the predicted benchmark rate of around 50% for the high confidence and 60%-75% for the medium confidence model when coding survey data [25]. This may be indicative of the quality of the data provided by UNISON survey respondents rather than the performance of the NIOCCS autocoding, however, data acquisition from patients at home versus health care settings should not vary on that basis alone. Because of this, the findings of this study may indicate the need to develop better methods, training, or emphasis on the collection of more detailed information on I&O to support successful autocoding.

## Limitations

Because the same 4-digit occupational code can exist within different 2-digit industry code groups, there is a potential for misinterpreting 4 digit codes from different 2-digit major categories as an exact match, however, it is unlikely that this occurred at a frequency high enough to bias the results: the 2-digit codes are quite different, and if it had happened at a frequency high enough to have biased the results, then the 4-digit agreement between hand coding and NIOCCS would have been inflated compared to the 2-digit agreement, which they were not; they were essentially equivalent.

It should be noted that true reliability of the codes—both in the collection phase and in the coding phase—would best be evaluated by further questioning of respondents (patients) to assure that the given I&O are accurate descriptors of their actual industrial sectors and job titles. This would require a study that entails responses to the I&O questions, followed by more extensive questioning of the respondent. In addition, the ability to evaluate hazardous workplace exposure, consider risk, or
promote health in the workplace would require more detailed questioning of the worker by the health care provider.

**Conclusions**

This study provides important field testing for the NIOCCS system on data collected through an ACA-required community needs assessment carried out by a university health system; development of an autocoding system was recommended by the Institute of Medicine’s Committee on Occupational Information and Electronic Health Records. Our results showed that the NIOCCS accuracy performed near its expected benchmark levels for 2-Major Groups SOC coding (“industry”), but well below the expected benchmark for the 4-digit detailed occupation SOC coding level (“occupation”) [23]. In this study, NIOCCS production rates fell below the anticipated production rates for survey data. This study could serve as an important baseline performance measure for NIOCCS as NIOSH continually improves the system to specifically target autocoding and accuracy rates.

According to the UNISON group, “Information learned from UNISON Health will help us to understand the health needs of the diverse community served by the UI Health and to use this information to improve health care for those who need it the most” [10]. “Work” is an important, though often ignored, determinant of health. Knowledge of adult patients’ places of employment—both I&O—can serve to inform interventions to improve health and well-being on a population level. Utilization of the NIOCCS tool could aid both researchers and this particular health system in understanding the occupational makeup of the population within its service area by reducing the time and cost associated with manually coding I&O information. It could also aid in establishing uniformity of I&O codes contained within patients’ EHRs that could be used to inform physicians of a patient’s unique work history and risks of work related health conditions.

**Acknowledgments**

NIOSH U60/OH009850 and T42/OH008672 provided funds for investigator time; the Office of the Vice President for Health Affairs of the shared the I&O data.

**Conflicts of Interest**

None declared.

**References**

7. President’s CP. April. 2010. Reducing environmental cancer risk - What we can do now URL: http://deainfo.nci.nih.gov/advisory/pcp/annualReports/pcp08-09npt/PCP_Renport_08-09_S08.pdf [accessed 2015-12-03] [WebCite Cache ID 6dX88er2Z]


**Abbreviations**

ACA: Affordable Care Act

EHRs: electronic health records

I&O: industry and occupation

NIOCCS: National Industry and Occupation Computerized Coding System

NIOSH: US National Institute for Occupational Safety and Health

SOC: Standard Occupational Classification

UNISON: University of Illinois Survey on Neighborhood Health
Disruptive Innovation: Implementation of Electronic Consultations in a Veterans Affairs Health Care System

Gouri Gupte1*, PhD, MHA; Varsha Vimalananda2,3*, MD; Steven R Simon4,5,6,7*, MD, MPH; Katerina DeVito8*, MSc; Justice Clark5, BS; Jay D Orlander5,9*, MD, MPH

1School of Public Health, Department of Health Policy and Management, Boston University, Boston, MA, United States
2Center for Healthcare Organization and Implementation Research (CHOIR), Edith Nourse Rogers Memorial Veterans Affairs Medical Center, Bedford, MA, United States
3Section of Endocrinology, Diabetes and Metabolism, School of Medicine, Boston University, Boston, MA, United States
4VA Boston Healthcare System, Center for Healthcare Organization and Implementation Research (CHOIR), Boston, MA, United States
5VA Boston Healthcare System, Section of General Internal Medicine, Boston, MA, United States
6Brigham and Women’s Hospital, Division of General Internal Medicine and Primary Care, Boston, MA, United States
7Harvard Medical School, Boston, MA, United States
8Harvard School of Public Health, Department of Nutrition, Boston, MA, United States
9School of Medicine, Section of General Internal Medicine, Boston University, Boston, MA, United States
* these authors contributed equally

Corresponding Author:
Gouri Gupte, PhD, MHA
School of Public Health
Department of Health Policy and Management
Boston University
Room 264
715 Albany Street
Boston, MA, 02118
United States
Phone: 1 617 414 1426
Fax: 1 617 638 5374
Email: gourig@bu.edu

Abstract

Background: Electronic consultations (e-consults) offer rapid access to specialist input without the need for a patient visit. E-consult implementation began in 2011 at VA Boston Healthcare System (VABHS). By early 2013, e-consults were available for all clinical services. In this implementation, the requesting clinician selects the desired consultation within the electronic health record (EHR) ordering menu, which creates an electronic form that is pre-populated with patient demographic information and allows free-text entry of the reason for consult. This triggers a message to the requesting clinician and requested specialty, thereby enabling bidirectional clinician-clinician communication.

Objective: The aim of this study is to examine the utilization of e-consults in a large Veterans Affairs (VA) health care system.

Methods: Data from the electronic health record was used to measure frequency of e-consult use by provider type (physician or nurse practitioner (NP) and/or physician assistant), and by the requesting and responding specialty from January 2012 to December 2013. We conducted chart reviews for a purposive sample of e-consults and semi-structured interviews with a purposive sample of clinicians and hospital leaders to better characterize the process, challenges, and usability of e-consults.

Results: A total of 7097 e-consults were identified, 1998 from 2012 and 5099 from 2013. More than one quarter (27.56%, 1956/7097) of the e-consult requests originated from VA facilities in New England other than VABHS and were excluded from subsequent analysis. Within the VABHS e-consults (72.44%, 5141/7097), variability in frequency and use of e-consults across provider types and specialties was found. A total of 64 NPs requested 2407 e-consults (median 12.5, range 1-415). In contrast, 448 physicians (including residents and fellows) requested 2349 e-consults (median 2, range 1-116). More than one third (37.35%, 1920/5141) of e-consults were sent from primary care to specialists. While most e-consults reflected a request for specialist input to a generalist’s question in diagnosis or management in the ambulatory setting, we identified creative uses of e-consults, including
requests for face-to-face appointments and documentation of pre-operative chart reviews; moreover, 7.00% (360/5141) of the e-consults originated from our sub-acute and chronic care inpatient units. In interviews, requesting providers reported high utility and usability. Specialists recognized the value of e-consults but expressed concerns about additional workload.

Conclusions: The e-consult mechanism is frequently utilized for its initial intended purpose. It has also been adopted for unexpected clinical and administrative uses, developing into a “disruptive innovation” and highlighting existing gaps in mechanisms for provider communication. Further investigation is needed to characterize optimal utilization of e-consults within specialty and the medical center, and what features of the e-consult program, other than volume, represent valid measures of access and quality care.


KEYWORDS
remote consultations; clinical communication; electronic consultation; telehealth; clinical information; decision making, telemonitoring; eHealth infrastructures

Introduction

Electronic consultations (e-consults) can be broadly defined as a platform for provider-provider consultation facilitated by health information technology such as an electronic health record (EHR) or Web-based portal. E-consults are most frequently used by primary care providers (PCPs) to access specialist input on clinical questions that can be addressed through chart review, thereby avoiding the need for a patient visit to the specialty clinic [1-3]. Generally, the intent of e-consults is to improve efficiency and overall quality of care by increasing access to specialty input while reducing unnecessary face-to-face visits [4-8].

While e-consults are a promising innovation, adoption is not widespread [3]. However, with health care reform in the United States and beyond [9], models of integrated care delivery may incentivize institutions to develop e-consult programs to increase access to specialty care while limiting costs [10]. Successful implementation will require a better understanding of how e-consults are being used and the reasons for those uses among both requesters, who are typically PCPs, and responders, who are typically specialists [3,4,11]. Ultimately, patients, clinicians, and policy-makers will need to know how e-consults influence the quality, safety, and cost of health care.

E-consults are a key component of the US Veterans Affairs (VA) health care system’s efforts to improve access to specialty care for Veterans, many of whom need to travel long distances to see their providers [8,12-16]. To gain insight into the ongoing rollout and implementation of e-consults across VA, we studied e-consult usage in a large VA health care system. We used quantitative and qualitative methods to characterize the various ways e-consults are used and to explore clinicians’ and hospital leaders’ attitudes and motivations underlying the patterns observed.

Methods

Study Design

An observational quality improvement study using mixed-methods was conducted. We extracted data from the EHR to quantify e-consult usage over time, by provider type (physician or nurse practitioner (NP) and/or physician assistant), and by specialty. We then used qualitative methods to explore both common and unexpected uses of e-consults. Our study utilized the mixed-methods sequential explanatory design described by Creswell and others [17,18]. This design involves collecting and analyzing first quantitative and then qualitative data in two consecutive phases within one study. The second phase builds on the first, in that the qualitative data and analysis help to explain the quantitative results obtained earlier. In our study, we conducted quantitative data collection and analysis to describe the patterns of e-consults use. We identified wide variation in uptake among users, as well as several unexpected patterns of use between and within services. We then used two qualitative methods to help explain our quantitative findings. We purposefully selected both high and low e-consult users as well as administrators, and interviewed them to better understand the reasons for variation in uptake across individuals, and to understand reasons for unexpected uses.

The primary analysis was conducted by three team members (GG, VV, and JC) and was reviewed with the entire study team. This analysis by the project team was cycled back and forth between individual cases and comparisons across cases to capture evolving themes and to understand the dynamics among users of e-consults. Throughout this process, the project team revisited the full interview notes for more detailed analysis on points of interest and to pursue hypotheses.

Setting

VA Boston Healthcare System (VABHS) is a tertiary care system consisting of three main campuses and five community-based outpatient clinics. VABHS provides primary care through a patient-centered medical home model to over 32,000 Veterans with more than 750,000 outpatient visits annually. The main campuses provide acute inpatient care, transitional, palliative, hospice, and nursing-home levels of care, as well as an extensive portfolio of specialized ambulatory procedures and consultative specialty clinics. VABHS is an academic center with multiple affiliations with health care training institutions, hosting hundreds of students, residents, and fellows each year. VABHS is the main referral center for five New England states.

Implementation of e-Consults at VABHS

In January 2011, e-consults were launched at VABHS for selected medical specialties [8,15]. The Department of Medicine
promoted use through discussions with primary care leadership
and emails to PCPs describing the availability and purpose of
e-consults. By early 2013, e-consults had expanded to all clinical
services, including all surgical specialties, mental health, and
pharmacy with the exception of radiology. Teledermatology
was not included as it has been used at VABHS for utilizing
collection, storing, and forwarding of new images, and is
considered another type of health information technology
application. Clinicians are able to request an e-consult in the
same way they request a face-to-face consultation, by requesting
the service from a menu within the EHR. Certain specialties
also accept e-consults from other VA facilities in New England.

Although e-consults in most health care systems are intended
for PCPs to request specialty consultations [3,10,19], the VA
platform allows any provider with ordering privileges to request
an e-consult [20]. The clinician selects the desired consultation
within the EHR ordering menu, which creates an electronic
form that is pre-populated with patient demographic information
and allows free-text entry of the reason for consultation. The EHR
has a built-in function that allows staff and clinicians to add
comments in a separate free-text field within the consultation
request. This triggers a message to the requesting clinician and
requested specialty, thereby enabling bidirectional
clinician-clinician communication. Each specialty has the ability
to tailor the e-consult form to solicit or require specific
information elements from the ordering clinician. For example,
some surgical specialties require the ordering clinician to
indicate whether the patient is taking anti-coagulant medications,
which would need to be considered before a surgical procedure.
Once the ordering clinician completes and electronically signs
the request for consultation, it is delivered electronically to the
requested specialty. Each specialty routes the incoming
consultation requests according to its own preferences. For example,
some specialties have a clerk perform an initial review
of all consultation requests before forwarding them to individual
specialist physicians, while others have one or more clinicians
receiving all the consultation requests directly from the ordering
clinicians. Consultants can choose to convert an e-consult to a
request for a face-to-face visit, or vice versa [21]. Responding
providers receive workload credit, a measurement of clinician
work that is used within VA for resource allocation, based upon
relative value units for each e-consult completed. In February
2014, VA approved the allocation of workload credit to
e-consults based on self-reported time spent completing the
e-consult in three discrete ranges (1) less than 15 minutes; (2)
15-30 minutes; and (3) greater than 30 minutes. These quantities
correspond to so-called levels 2, 3, and 4 for outpatient visit
complexity. Responding providers are expected to answer
e-consults within three working days.

Data Collection and Analysis
We used the VA EHR to extract quantitative data from January
2012 to December 2013. Data included information on the
sending and receiving provider type and specialty, and the date
and time of actions on each e-consult. All specialties except
radiology participated in the usage of e-consults at the point of
the data collection. The date and time stamps allowed us to
calculate time to completion, measured as the elapsed time
between the signature of the requesting clinician on the e-consult
request and the signature of the clinician providing consultation
on the completed e-consult form. We used findings from the
quantitative data to inform 30-minute, semi-structured interviews
with frequent and less frequent requesters and responders and
both clinical specialty and hospital leadership in order to better
understand e-consult utilization as well as barriers and
facilitators of use. We contacted and interviewed 17 medical
doctors (MDs), 9 NPs, 1 doctor of pharmacy, and 4 hospital
leaders (including 2 chiefs of specialties) for a total of 31
providers from 21 specialties. All individuals we contacted
agreed to participate. Using a semi-structured interview guide,
we asked providers about their knowledge, usage, experience,
and feedback related to e-consults. We asked leadership about
the strategies used for promoting e-consult implementation and
uptake, assessment methods, and future plans. Interviews were
audio recorded and transcribed verbatim for analysis. We held
team meetings to identify thematic categories related to our
project goals, refine their meaning, discuss alternative
interpretations, and reach an agreement on representative
quotations for each category.

This work was reviewed by the VABHS Institutional Review
Board and was determined to be quality improvement rather
than human subjects’ research.

Results
Quantitative Findings
Our dataset contained information on a total of 7097 e-consults,
representing all VABHS e-consults during the study period. In
2012, 1998 (28.15%, 1998/7097) e-consults were completed,
compared with 5099 (71.85%, 5099/7097) in 2013, representing
a 150% year-to-year increase. More than one quarter (27.56%,
1956/7097) of the e-consult requests originated from VA
federations in New England other than VABHS. Further analyses
were limited to the 5141 e-consults originating within VABHS.
The distribution of clinical locations from where e-consults
originated within VABHS are displayed in Table 1. More than
one-third of the requests for e-consults (37.35%, 1920/5141)
originated in primary care, representing the single largest
requesting specialty. Nearly one-third (32.0%, 1645/5141) of the
e-consults were directed to general surgery and surgical
subspecialties.
<table>
<thead>
<tr>
<th>Specialties</th>
<th>Number of e-consults sent, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medicine</strong></td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>1920 (37.35)</td>
</tr>
<tr>
<td>Other medical sub-specialties&lt;sup&gt;a&lt;/sup&gt;</td>
<td>398 (7.74)</td>
</tr>
<tr>
<td><strong>Surgery</strong></td>
<td></td>
</tr>
<tr>
<td>Orthopedics</td>
<td>634 (12.33)</td>
</tr>
<tr>
<td>Pre-admission testing clinic</td>
<td>370 (7.20)</td>
</tr>
<tr>
<td>General surgery</td>
<td>211 (4.10)</td>
</tr>
<tr>
<td>Other surgical sub-specialties&lt;sup&gt;b&lt;/sup&gt;</td>
<td>542 (10.54)</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
<tr>
<td>Sub-acute chronic and inpatient care</td>
<td>365 (7.10)</td>
</tr>
<tr>
<td>Medical acute care</td>
<td>136 (2.65)</td>
</tr>
<tr>
<td>Mental health/psychiatry</td>
<td>75 (1.46)</td>
</tr>
<tr>
<td>Other sub-specialties&lt;sup&gt;c&lt;/sup&gt;</td>
<td>490 (9.53)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Medical sub-specialties include pulmonary (2.12%, 109/5141), cardiology (1.44%, 74/5141), gastroenterology (1.09%, 56/5141), geriatrics (0.70%, 36/5141), renal (0.58%, 30/5141), rheumatology (0.49%, 25/5141), oncology (0.35%, 18/5141), sleep (0.31%, 16/5141), post-discharge clinic (0.23%, 12/5141), hematology (0.18%, 9/5141), infectious disease (0.12%, 6/5141), dermatology (0.06%, 3/5141), allergy (0.04%, 2/5141), endocrinology (0.02%, 1/5141), and palliative care (0.02%, 1/5141).

<sup>b</sup>Surgical sub-specialties include optometry (2.49%, 128/5141), ear nose and throat (2.22%, 114/5141), urology (1.50%, 77/5141), ophthalmology (1.09%, 56/5141), thoracic (0.97%, 50/5141), vascular (0.41%, 21/5141), gynecology (0.31%, 16/5141), podiatry (0.21%, 11/5141), bariatric (0.11%, 6/5141), cardiac surgery (0.08%, 4/5141), and plastic surgery (0.06%, 3/5141).

<sup>c</sup>Other includes administrative (2.82%, 145/5141), undefined (1.24%, 64/5141), neurology (1.19%, 61/5141), urgent care (1.46%, 75/5141), spinal cord injury (1.09%, 56/5141), radiology (0.78%, 40/5141), pharmacy (0.62%, 32/5141), surgical acute care (0.49%, 25/5141), rehab medicine (0.31%, 16/5141), anesthesia (0.19%, 10/5141), occupational health (0.14%, 7/5141), audiology (0.08%, 4/5141), dental (0.12%, 6/5141), nutrition (0.04%, 2/5141), prosthetics (0.04%, 2/5141), and radiation therapy (0.02%, 1/5141).

The median time to completion across specialties was 2.2 working days (range 0.8-56). A subset of e-consults (45.77%, 2353/5141) included consultants’ indication of time spent completing the e-consult. Most (83.00%, 1953/2353) indicated that they spent less than 15 minutes, while 11.00% (259/2353) spent 15-30 minutes and 5.00% (141/2353) spent more than 30 minutes completing the e-consult. We identified variability in the number of e-consults requested by individual clinicians. A total of 64 NPs requested 2407 e-consults (median 12.5, range 1-415). In contrast, 448 physicians (including residents and fellows) requested 2349 e-consults (median 2, range 1-116). A total of 385 e-consults, representing 7.49% (385/5141) of all e-consults during the study period, were submitted by staff members other than NPs and physicians. Department of Medicine physicians requested a median of 2.0 e-consults (range 1-103), while NPs in Medicine requested a median of 5.0 e-consults (range 1-88).

The distribution of specialties at VABHS receiving the most e-consults requests is displayed in Figure 1. The medical specialties consulted most frequently were cardiology (13.73%, 706/5141), hematology (9.92%, 510/5141), sleep medicine (9.16%, 471/5141), gastroenterology (6.44%, 331/5141), and pulmonary (5.72%, 294/5141). Frequently consulted specialties within surgery included orthopedics (12.92%, 664/5141) and ophthalmology (2.57%, 132/5141). We unexpectedly identified a number of e-consults submitted within one specialty. For example, 27.72% (487/1757) of e-consults to surgery and surgical specialties were submitted by a clinician in orthopedics to the orthopedics service (ie, an intra-specialty e-consult). These intra-specialty e-consults to orthopedics accounted for 9.47% (487/5141) of all e-consults within the VABHS (see Figure 1). Similarly, 5.74% (101/1757) of e-consults to surgery and surgical specialties were submitted by optometry to ophthalmology, while 1.76% (31/1757) of e-consults to surgery and surgical specialties were submitted by ophthalmology to optometry.
Qualitative Findings

Unanticipated Uses of e-Consults
The flexibility of the e-consult mechanism was often considered an advantage by specialists, as each specialty could develop its own triage and response mechanism. This flexibility also allowed adaption of the e-consult mechanism for several uses that were not originally intended. Quantitative data had revealed a number of e-consults submitted within the same specialty, or within two divisions of the same specialty. For example, we identified numerous e-consults requested from optometry to ophthalmology and vice-versa. Upon inspection of their content and interviewing providers, we found the majority of these e-consults were submitted to facilitate appointment scheduling. Some were scheduling requests within the same specialty at VABHS, while others were requests for assistance in scheduling post-ophthalmologic follow-ups at VA sites closer to patients’ homes.

We identified 588 within-specialty e-consults, particularly among surgical specialties. Through chart review and interviews we found that the majority of these within-specialty e-consults, most frequently occurring in orthopedics, but also in general surgery and other surgical specialties, had been adapted to document and facilitate pre-operative assessments. We ascertained that NPs would generate an e-consult to themselves to document a detailed pre-operative chart review. Based upon this review, NPs often solicited consultation from other specialties through additional e-consults to address issues such as pre-operative evaluation of cardiac risk (“cardiac clearance”) and pre-operative medication management (eg, anticoagulation). This work occurred prior to scheduling the patient for the surgical procedure and prior to the NP obtaining a formal history and conducting a physical examination in the pre-operative testing clinic; clinicians explained that the intention was to decrease the likelihood that scheduled elective procedures would need to be delayed. Prior to implementation of the e-consult system, NPs documented their chart review in various locations of the medical record. The e-consult system provided a uniform and easy-to-find location for such documentation that also ensured documentation of the clinical work completed (ie, workload credit). Benefit to patients was noted, as one surgical subspecialty chief said, “I think they (patients) benefit by having all their care coordinated before they come down here, too, so they don’t have to keep coming back multiple times.”

Benefits of e-Consults From Primary Care Providers’ Perspective
PCPs that used e-consults frequently were unanimous in the opinion that e-consults were easy to use, useful, and increased timeliness of and access to specialty care. As one provider said, “I’m a huge fan of it and a big advocate.” PCPs reported that placing e-consults was straightforward and that they were “pretty easy to find in the EHR.” The free-text format of the e-consult allowed a welcome flexibility in writing the reason for consultation with greater or lesser amounts of detail, depending on the clinical situation. The usefulness of e-consults was related to both the rapidity of the specialist response and the provision of a mechanism for asking simple clinical questions. One PCP noted, “I would define an e-consult as a higher-level question for a specialist that could be safely answered by a chart review.” Having a form of communication that was clinically oriented but asynchronous in nature seemed to empower PCPs to seek formal specialty consultation in circumstances where they previously would seek consultation informally such as via a hallway conversation, an e-mail, or phone call, often labeled curbside consultations [20]. As a provider commented, “Things that historically would have been curb-sid ed can now be
documented as a formal conversation.” PCPs also noted that they now found themselves completing e-consult requests in situations where they previously might not have pursued specialist input due to barriers in reaching a specialist colleague and when formal referral of the patient for face-to-face visit did not seem indicated.

The easy availability of e-consults was particularly valuable for primary care NPs at community-based outpatient clinics. For these providers, the geographic distance between themselves and the hospital-based specialists generally precluded access to specialist input through informal conversations. Most PCPs perceived that e-consults resulted in fewer face-to-face consultations, more efficient medication management, expedited diagnostic testing in lieu of or in preparation for a specialty visit, and more effective communication with specialists. We also interviewed two providers who infrequently used e-consults who stated their reason for low use was because they were unfamiliar with the process itself.

**Benefits of e-Consults From Specialists’ Perspective**

Like PCPs, specialists noted a major benefit of e-consults to be fewer unnecessary face-to-face consultations. Specialists perceived more available appointment times for scheduling patients who did require face-to-face visits. As one specialist said, “It saves patients a lot of time, it makes us more efficient because we can take care of the patients who really need our services and are really sick, and, ultimately, you can probably reduce manpower too.” Specialists particularly valued having the ability to convert face-to-face consultation requests to e-consults when they deemed it appropriate.

Converting face-to-face consultation requests to e-consults was most prevalent in sleep medicine. Sleep medicine physicians, faced with extremely long wait times for clinic appointments, found that in many cases chart review was sufficient to identify which patients were at high risk for having obstructive sleep apnea, and therefore required a sleep study. Conversion of face-to-face consultation requests to e-consults allowed the sleep medicine specialist to render an opinion and order a sleep study for the patient within a few days without the need for a face-to-face evaluation. Sleep medicine physicians indicated that this approach was reported to improved clinic access and dramatically reduced time to ordering and completing a sleep study. In many cases, patients would be diagnosed and started on therapy before having their first live face-to-face consultation visit in the sleep medicine clinic. Sleep medicine physicians did question whether the quality of care and adherence to treatment would be improved if patients had been seen by them before diagnostic testing was completed and therapy initiated.

Specialists discussed the potential for e-consults to support PCP education, in that the consult questions often represented gaps in clinical knowledge that could be addressed with detailed information that might later be used as a reference. They noted that success of this strategy depended on the engagement of the referring provider in the learning process and that some clinicians made nearly identical referrals repeatedly. This pattern was observed most commonly among e-consults submitted by NPs in surgical specialties to medical subspecialties, such as hematology and cardiology, as part of a pre-operative assessment.

**Concerns About e-Consult Implementation and Impact on Workflow**

While PCPs did not report concerns about e-consults, specialists receiving the highest number of e-consults reported workflow and workload issues due to the increase in volume overtime. Specifically, they complained about the lack of a pre-implementation evaluation of e-consults’ impact on the work of the section specialty and the individual providers. E-consults originating outside of VABHS were particularly challenging and time-consuming to complete, since the mechanism for reviewing relevant information outside the main facility involves use of an associated EHR software package that is not easily searchable for specific data. Satisfaction with e-consults varied greatly between individual specialists. For example, a hematologist felt his specialty was particularly well suited to e-consultation and estimated that the hematology unit completed nearly 50% of all new requests for consultations via the e-consult process. In contrast, another individual specialist described e-consults as “an unfunded mandate” created by leadership, which was poorly integrated within the existing workload and fraught with potential legal concerns based on providing advice on a patient who was never examined in person.

One specialist reported that e-consults could contribute to breaks in continuity of care, as the respondent to an e-consult may be different from the specialist who previously saw a given patient in person. Many specialists noted that similar questions were asked repeatedly by the same providers, suggesting to them that the purpose of the e-consult for some individuals was documentation rather than clinical support and knowledge transfer. While they understood that some providers may want documentation due to a lack of clinical confidence, the e-consults for this purpose created excess and unwelcome work for some consultants.

**Leadership Engagement to Promote e-Consult Usage**

Interviews with hospital and department-level leadership revealed strong enthusiasm for e-consults. One clinical service chief noted “It allows for better triaging of consults and in theory may be improving access to subspecialty clinics.” Leaders promoted e-consults in a manner similar to how they encouraged their constituencies to embrace other information technology initiatives to improve specialty access, such as telehealth. Under this model, service-level leadership delegated staff to act as champions for e-consults across clinical services. Two hospital leaders discussed the challenges of implementing e-consults, noting “There was a lot of infrastructure building for e-consults and a lot of education that had to be done.” They described a need for more effective strategies to encourage uptake across providers and the potential benefit of incorporating structured fields with mandatory data elements in the e-consult forms to ensure that relevant information is included both by providers requesting consultation and those providing it.

Hospital leaders were not explicitly involved in measuring or tracking the use of e-consults. Use across the medical center

http://medinform.jmir.org/2016/1/e6/
was monitored at the regional Veterans Integrated Service Network (VISN) level only at the time of our interviews. They understood that the number of e-consults is tracked by specialty, with and considered a proxy for access to specialist expertise. VABHS hospital administrators did not have an available mechanism to provide a more detailed assessment of the content, quality or impact of e-consults on an ongoing basis.  

Recommendations From Users of e-Consults  
We closed our semi-structured interviews by asking for recommendations to improve the e-consult process. PCPs suggested that leadership should reach out to more providers, specialties, and specialists to improve buy-in of the e-consult process and increase both the number of clinicians requesting e-consults and the number of specialties providing consultation by this mechanism. Specialists suggested specialty-specific discussions to address the perception of an “unfunded mandate,” individual performance measurement, legal issues, workflow consideration, and measurement of outcome metrics by the usage of e-consults to create more awareness and engagement. One particular request made during the interviews was about clarity on how credit for work completed is assigned for consultants, as only the lowest level of workload credit was allowed when e-consults were initially launched. VA addressed this in February 2013 with nationwide approval of three different levels of workload credit based upon self-reported time spent completing the e-consult (reported in the quantitative results above). Leadership discussed the need for better evaluation strategies, which could include patient satisfaction scores, provider feedback, process measures, and qualitative workflow impacts.

Discussion  
Principal Findings  
We undertook this quality-focused evaluation to characterize the usage of e-consults in a large VA health care system, and to describe how the experience of providers and hospital leaders shapes such use. Nearly one-third of e-consults originated from clinicians outside VABHS, representing robust use of the e-consult mechanism to improve specialty access for patients at locations with limited specialty services. As expected, PCPs were the most frequent requesters of e-consults. Using qualitative methods to better understand patterns observed in our quantitative data, we identified innovative and unexpected ways that e-consults are being used to expedite evaluation of sleep disorders, for administrative communication between related specialties, and for pre-operative documentation within a specialty. Overall, both PCPs and specialists felt that e-consults improved timeliness of specialty input and reduced unnecessary face-to-face consultations. PCPs were generally highly satisfied with e-consults, while specialists had concerns related to workload and workflow.

As reported in other health care systems, we found that cardiology and hematology were frequent recipients of e-consults [19,22,23]. Further study is needed to determine whether this finding represents an overall high volume of patient need for such services, particular suitability of e-consults to these specialties, or a combination of these and other factors. The particularly high use of e-consults in sleep medicine and orthopedics was largely explained by the specific use of e-consults as adaptations to other system shortcomings; in the former, limited specialty clinic availability, and in the latter, a lack of means to document clinical work that did not include a patient visit. As noted in other studies, dermatology e-consults were not observed in the data as they present as “teledermatology” utilizing collection, storing, and forwarding of new images, and are considered another type of health information technology application [5]. Many of the e-consults to ophthalmology, the second most common recipient in surgical specialties, were requests from optometry for assistance in arranging logistically complicated post-operative plans. The adoption of e-consults for uses other than the originally intended purpose suggests that e-consults in our health care system have become a “disruptive innovation.”

Disruptive usage begins when an innovative product takes root among a group of users because it addresses previously unmet needs [24-27]. As knowledge spreads about the utility of the innovation to address these unmet needs, the innovation rapidly garners support and utilization climbs [25]. Disruptive innovations are generally convenient, easy to use, and simpler than existing or prior systems [25,27]. In addition, they are aligned and blend with a pre-existing technology infrastructure; in the case of e-consults, the existing infrastructure is the EHR. E-consults have emerged within VA as a disruptive innovation in part because the existing infrastructure and clinical processes were sufficiently flexible to tolerate the innovation. The VA EHR allows any provider with ordering privileges to request an e-consult from any other service; any type of question or request could be entered into the free-text field, without constraint to specific clinical situations or diagnoses. Hence, the availability of e-consults has led to providers asking more questions of their specialty colleagues. Furthermore, each specialty has the flexibility to develop its own mode of triage and assignment of responsibility for responding to submitted e-consults.

We found a rapid increase in e-consult use since 2011, with high levels of satisfaction among requesters. Users have actively been promoting e-consult use amongst their peers, with the encouragement of clinical leadership. These users are applying e-consults in unexpected ways to address various shortcomings with existing processes; in short, the avid users are making e-consults a disruptive innovation. One result of disruptive innovation is that it identifies previously hidden needs in the current system [24-26]. Our work raises the question of whether unexpected uses of e-consults as workarounds for system needs represent appropriate use. Unexpected and unintended uses of e-consults should be explored in other health care settings, and as e-consult use expands, each health care system will need to develop protocols and policies for managing such unanticipated use and its consequences.

High satisfaction with e-consults among PCPs has been reported in studies of e-consult use in other health care systems. Consistent with those studies, PCPs in this study appreciated the ease of access to specialists, the timeliness of specialty input, and the perception of reduced travel for patients [3,10,28]. In comparison, we identified considerable variation in satisfaction.
among specialists, a finding that is also consistent with prior work [4,11,20,29]. While most specialists perceived a reduction in unnecessary face-to-face visits, some were concerned about overall higher workload due to increasing e-consult volume. Addressing specialists’ concerns about the use of e-consults will be essential as this technology is expanded more broadly. The utility of e-consultation and ultimately the success of the e-consult program will depend in part on the willingness of specialists to answer appropriate e-consult requests as presented, rather than converting them to face-to-face consultations, and on the quality of the specialists’ response itself. How specialists accept and respond to e-consults are, to some degree, functions of the types of questions being asked by requesters and by the perceived benefits of e-consultation to the specialist. Our work suggests that better communication between specialists, requesting providers, and leadership about the needs and expectations of each group would foster increased uptake among specialists. While not addressed directly in our interviews, the role of the patient in e-consults merits further examination.

Concurrent with our work, some tangible improvements in the use of e-consults occurred in our institution. During our interviews, we found that one specialty was unaware of the mechanisms for identifying incomplete consultations (ie, those consultations that were requested but unanswered). Through other interviews, we identified problems in routing of consultations to incorrect staff in another service. In both cases, we provided a brief tutorial, improving the specialty’s management of their e-consult load. We also provided feedback to our office of information systems on the menu within the EHR for selecting e-consults; this feedback led to reorganization of the types of e-consults in the menu, improving usability. While this study was not designed for rigorous evaluation of the effects of e-consults on clinical practice or outcomes. Our findings have been received with interest among primary care and specialist clinical leaders within our institution and in the regional network of VA facilities.

Limitations
We undertook this project as a quality improvement study in a single health care system within VA. Though the specifics of e-consult use may vary across sites, PCP-to-specialist use is a common finding. Numerous specialties were represented in our data, but in most of them there were a small number of specialist physicians actually involved in triaging and responding to e-consults, limiting the generalizability of our findings. However, our mixed-methods approach of starting with qualitative data, examining de-identified e-consults from frequent users, and following up with semi-structured interviews helped us better understand use, workflow, and variability. Examining data more broadly across VA and eventually to other health care systems would provide a more robust characterization of all these issues. Another possible limitation of our analysis is that we used an administrative database, which did not include certain information about e-consults, such as whether they were converted to or from face-to-face consultations, or the time to answer. Thus, our count of the numbers of e-consults may underestimate the true quantity. However, since our aim was to identify trends in e-consult use, rather than absolute quantity, we do not believe inclusion of these e-consults would meaningfully change the trends we observed. Our analyses excluded more than one-fourth of the e-consults that were submitted from outside VABHS, as full information on these was not available at the time of the study. Future evaluations should explore the content and circumstances of these consultations in comparison with those submitted within VABHS.

E-consults generate a modest amount of workload credit for those responding to the request for consultation, but this workload credit is an internal metric and not linked to insurance claims or any other external monitoring system. As such, there is no built-in audit system in place to assure quality or appropriateness of use. Our findings strongly suggest that the number of e-consults is not a valid metric for access to specialty care services, such that reliance on measures of volume is likely to result in erroneous conclusions about access to care. Specific examination of the clinical content of e-consults may be necessary to determine the extent to which this innovation is actually improving Veterans’ access to high quality specialty care, a major goal of improving overall care at the VA [30].

Conclusions
Our study shows that e-consults facilitate access to specialty expertise for PCPs, but that use across clinical services has become more widespread than originally intended. Further investigation across other VA systems is warranted in order to identify best practices as well as pitfalls of the e-consult mechanism. Additional work is needed to define what features of an e-consult program represent valid measures of access and quality care, and what monitoring systems, if any, need to be implemented.

Conflicts of Interest
None declared.

References


Abbreviations

e-consult: electronic consultation
EHR: electronic health record
PCPs: primary care providers
NP: nurse practitioner
VA: Veterans Affairs
VABHS: Veterans Affairs Boston Healthcare System

©Gouri Gupte, Varsha Vimalananda, Steven R Simon, Katerina DeVito, Justice Clark, Jay D Orlander. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 12.02.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.
Improving Inpatient Surveys: Web-Based Computer Adaptive Testing Accessed via Mobile Phone QR Codes

Tsair-Wei Chien¹,², MBA; Weir-Sen Lin², MPH, PhD
¹Chi Mei Medical Center, Taiwan, Research Department, Chi Mei Medical Center, Taiwan, Tainan, Taiwan
²Chia-Nan University of Pharmacy and Science, Department of Hospital and Health Care Administration, Chia-Nan University of Pharmacy and Science, Tainan, Taiwan

Corresponding Author:
Weir-Sen Lin, MPH, PhD
Chia-Nan University of Pharmacy and Science
Department of Hospital and Health Care Administration
Number 60, Section 1, Erren Road, Rende District
Tainan City 71710
Taiwan (R.O.C.)
Phone: 886 +886 6 2664911
Fax: 886 +886 6 2664911
Email: weirsen@mail.cnu.edu.tw

Abstract

Background: The National Health Service (NHS) 70-item inpatient questionnaire surveys inpatients on their perceptions of their hospitalization experience. However, it imposes more burden on the patient than other similar surveys. The literature shows that computerized adaptive testing (CAT) based on item response theory can help shorten the item length of a questionnaire without compromising its precision.

Objective: Our aim was to investigate whether CAT can be (1) efficient with item reduction and (2) used with quick response (QR) codes scanned by mobile phones.

Methods: After downloading the 2008 inpatient survey data from the Picker Institute Europe website and analyzing the difficulties of this 70-item questionnaire, we used an author-made Excel program using the Rasch partial credit model to simulate 1000 patients’ true scores followed by a standard normal distribution. The CAT was compared to two other scenarios of answering all items (AAI) and the randomized selection method (RSM), as we investigated item length (efficiency) and measurement accuracy. The author-made Web-based CAT program for gathering patient feedback was effectively accessed from mobile phones by scanning the QR code.

Results: We found that the CAT can be more efficient for patients answering questions (ie, fewer items to respond to) than either AAI or RSM without compromising its measurement accuracy. A Web-based CAT inpatient survey accessed by scanning a QR code on a mobile phone was viable for gathering inpatient satisfaction responses.

Conclusions: With advances in technology, patients can now be offered alternatives for providing feedback about hospitalization satisfaction. This Web-based CAT is a possible option in health care settings for reducing the number of survey items, as well as offering an innovative QR code access.

(JMIR Med Inform 2016;4(1):e8) doi:10.2196/medinform.4313

KEYWORDS
computer adaptive testing; patients’ experience; partial credit model; consultation experience and perception; smart phone

Introduction

Patient reports are central to the evaluation of medical care, both in terms of treatment outcomes (ie, patient-reported outcomes and in terms of experiences of quality of care (ie, patient-reported experience measures) [1]. A quality standard for patient experience in the United Kingdom’s National Health Service (NHS) has been developed by the National Institute for Health and Care Excellence [2,3]. The UK National Adult Inpatient Survey—UK NHS 70-item questionnaire—has been in use in Great Britain since 2002, gathering data from over 620,000 patients every year [4].
The first public reporting of the US equivalent of such surveys, the Hospital Consumer Assessment of Healthcare Provider and Systems (HCAHPS), began in 2008. HCAHPS collects more than 3.0 million completed surveys from 3912 hospitals every year. On average, more than 28,000 patients are surveyed every day about their recent hospital experience, and more than 8400 patients (approximately 30% response rate) complete the HCAHPS inpatient survey every day [5].

Research Questions

A concern reported in the literature is the burden on patients of answering all survey questions at one time [6-9]. More than 3.6 million patients completed mail-in surveys from the UK NHS (600,000 patients). The US survey (3.0 million respondents) required approximately 6 hours of time per month and cost approximately US $70 per month [3] to examine people’s health service experiences [10]. However, the UK NHS 70-item questionnaire is significantly longer than the US HCAHPS 25-item survey [11,12]. To reduce patient burden, it is first necessary to shorten the item length of the UK NHS inpatient questionnaire to increase response rates without compromising its assessment reliability [2,13].

Many studies [6-9] have reported that item response theory (IRT)-based computer adaptive testing (CAT) has the advantages of both long-form and short-form questionnaires [14-16] in precision and efficiency. Since many patients (or their guardians) already own mobile phones, which they are comfortable using, it makes sense to use them in hospitals and for hospital surveys. At this time, no studies have been published reporting online CAT via mobile phones in medical fields.

However, many skip items (see Multimedia Appendices 1 and 2) exist in the UK NHS 70-item questionnaire, which can be confusing and may perplex researchers on CAT implementation. Thus, our second aim was to tackle the problem of skip items in the UK NHS questionnaire and to implement the online CAT.

Rasch Partial Credit Model Applied to the Item Response Theory-Based Computer Adaptive Testing

Many researchers have contributed to the dichotomous [6] and polytomous [7-9] formats used by CAT. The UK NHS questionnaire comprises items with different categories (eg, 3 and 6 categories for Items 40 and 41; see Multimedia Appendix 1). It is suited for applying the Rasch partial credit model (PCM), that is, items with a different number of responses and with an equal discrimination parameter [17], or the generalized partial credit model, that is, items with a different number of responses and with unequal discrimination parameters [18], if those items form a unidimensional construct. None was jointly available for a comparison of precision and efficiency differences of CAT estimation with the aforementioned methods commonly used in literature, such as PCM, answering all items (AAI), and the randomized selection method (RSM).

Further, as mobile phones have become ubiquitous in the health care setting [19], it is important to offer an alternative online Rasch PCM-CAT assessment to gather hospitalization experience feedback from patients. We propose access to the questionnaire using a quick response (QR) code via mobile phone.

Aims of this Study

The aims of the current study were to investigate whether CAT can (1) be efficient with item reduction and (2) be used with QR codes used for mobile phones.

Methods

Study Data

The UK NHS 70-item questionnaire regarding patient experience was downloaded from the NHS official website [11]. The item and its threshold difficulties (lower summation scores for an item imply that it was more difficult for examinees to respond) were roughly determined by hand computation according to the key findings report for the 2008 inpatient survey [20]. We simulated an interactive metric of 1000 persons (following a normal distribution ~N(0,1)), called true scores) and 70 items (estimated with aforementioned item difficulties) using the Rasch PCM model [17,21]. Nine items originally designed to automatically select different paths were set with different probabilities by the authors. The remaining 61 items were allocated different weighted scores (see Multimedia Appendix 1). A set of 24 items (ie, regarding sections of the ward, doctors, nurses, patient care, and treatment) was extracted from the UK NHS 70-item questionnaire to be the CAT item pool (see Multimedia Appendix 2). We assumed that the set of 24 items is unidimensional based on the report from the previous study paper [13]. Because these 9 conditional selection path items make CAT difficult to design for a computer, they were excluded from the CAT item pool. Multimedia Appendix 3 shows the file layout and fields we designed for use with the datasets.

Unidimensionality

The Rasch model, named after Georg Rasch [22], is a psychometric model for analyzing categorical data as a mathematical function of the trade-off interaction between (1) the respondent’s latent trait (eg, hospitalization perception level in this study) and (2) the item difficulties. The dichotomous Rasch model and its extensions (eg, family models: rating scale model [23], PCM [17]) are successfully used in other areas, including the health profession [24] and market research [25], because of their general applicability [26].

The study data need to meet the following criteria to fit the Rasch model: the infit and outfit mean square errors (MNSQ) of all items are <1.5 for unidimensionality and >0.5 for local independence [27]. Simulation data were generated fitting to the Rasch PCM model [21].

Task 1: Investigating Computer Adaptive Testing Efficiency and Accuracy

Three scenarios were designed to compare their efficiency and accuracy on the UK NHS 70-item questionnaire: (1) the AAI (answering all items on those 24 items), (2) the RSM (randomized selection method to draw 12 items), and (3) the CAT (at least 5 items and stop at person reliability of 0.80) responding to the 24-item pool.

We applied CAT stop rules, such as when person reliability reaches 0.80 (\[1 − SEM_{pi}\]×[1 − SEM_{pi}]), where SEM_{pi}=person
standard error of measurement on item $i=1/variance_{pi}=1/information_{pi}$, and when the last 5 average consecutive person estimation change is less than 0.05 after the minimum necessary completed number of items is ≥5.

In addition, we ran an author-made VBA (Visual Basic for Applications) module in Microsoft Excel to conduct a simulation study (see Multimedia Appendix 4). Another Web-based CAT was programmed for use on mobile phones. The maximum likelihood estimation algorithm [28] (see Multimedia Appendix 4) was used to (1) estimate person measures on the three scenarios, (2) compute correlation coefficients between estimated person measures among the three scenarios and the original true scores to verify CAT accuracy, and (3) analyze CAT efficiency of item length shortened by CAT compared with the other two scenarios (ie, AAI and RSM) (see Figure 1).

**Figure 1.** Study flowchart.

---

**Task 2: An Online Assessment Using Mobile Phones**

**The Conditional Path Skip Items Designed on UK NHS-70**

The path skip item was automatically redirected to the next according to the respective route designed in the field of the item dataset (see Multimedia Appendix 3). To illustrate Item 39 in Multimedia Appendix 1, two route fields were filled with Items 40 and 41 in response to the respective answer code (eg, 1 or 2). In contrast, the route fields for those ordinary non-skip items were kept empty (or a null value).

**An Online Computer Adaptive Testing Routine for Gathering Feedback From Patients**

An online routine was designed for patients to report their perceptions of their inpatient hospitalization experience. The UK NHS 70-item questionnaire (see Multimedia Appendix 2) was uploaded to website. The first CAT item is randomly selected from the item pool (ie, Items 15-38) after Item 14 is answered. The next item to be answered is the item with the maximal variance in the remaining items according to the provisional person ability [7,29]. Multimedia Appendix 5 shows details on the item selection rules and the Excel VBA codes for the conditional path items. All the responses are then automatically saved on the study website (see Multimedia Appendix 3).

**Results**

**Items Fit to the Rasch Model**

The set of 24 items (see Table 1) was taken as unidimensional due to simulation data fitting the Rasch model’s requirement with values of infit and outfit MNSQ between 0.5 and 1.5 [21].
Each item has its own threshold difficulties (see Multimedia Appendix 4).

### Table 1. The 24 items selected from the UK NHS 70-item questionnaire.

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Threshold</th>
<th>Difficulty</th>
<th>In MNSQ</th>
<th>Out MNSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Were you ever bothered by noise at night from other patients?</td>
<td>1</td>
<td>2.17</td>
<td>0.96</td>
<td>1.00</td>
</tr>
<tr>
<td>16</td>
<td>Were you ever bothered by noise at night from hospital staff?</td>
<td>1</td>
<td>1.84</td>
<td>1.03</td>
<td>1.11</td>
</tr>
<tr>
<td>17</td>
<td>In your opinion, how clean was the hospital room or ward that you were in?</td>
<td>3</td>
<td>-0.56</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>18</td>
<td>How clean were the toilets and bathrooms that you used in the hospital?</td>
<td>3</td>
<td>-2.20</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>19</td>
<td>Did you feel threatened during your stay in the hospital by other patients or visitors?</td>
<td>1</td>
<td>1.88</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>20</td>
<td>Were hand-wash gels available for patients and visitors to use?</td>
<td>2</td>
<td>1.33</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>21</td>
<td>How would you rate the hospital food?</td>
<td>3</td>
<td>-1.72</td>
<td>0.96</td>
<td>1.02</td>
</tr>
<tr>
<td>22</td>
<td>Were you offered a choice of food?</td>
<td>2</td>
<td>-0.04</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>23</td>
<td>Did you get enough help from staff to eat your meals?</td>
<td>2</td>
<td>0.98</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>24</td>
<td>When you had important questions to ask a doctor, did you get answers that you could understand?</td>
<td>2</td>
<td>-1.01</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>25</td>
<td>Did you have confidence and trust in the doctors treating you?</td>
<td>2</td>
<td>-0.86</td>
<td>1.06</td>
<td>1.07</td>
</tr>
<tr>
<td>26</td>
<td>Did doctors talk in front of you as if you weren’t there?</td>
<td>2</td>
<td>-1.15</td>
<td>0.96</td>
<td>1.00</td>
</tr>
<tr>
<td>27</td>
<td>When you had important questions to ask a nurse, did you get answers that you could understand?</td>
<td>2</td>
<td>1.48</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>28</td>
<td>Did you have confidence and trust in the nurses treating you?</td>
<td>2</td>
<td>1.26</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>29</td>
<td>Did nurses talk in front of you as if you weren’t there?</td>
<td>2</td>
<td>-1.46</td>
<td>1.04</td>
<td>1.05</td>
</tr>
<tr>
<td>30</td>
<td>In your opinion, were there enough nurses on duty to care for you in the hospital?</td>
<td>2</td>
<td>-1.40</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td>31</td>
<td>Sometimes in a hospital, a member of staff will say one thing and another will say something quite different. Did this happen to you?</td>
<td>2</td>
<td>-1.22</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>32</td>
<td>Were you involved as much as you wanted to be in decisions about your care and treatment?</td>
<td>2</td>
<td>0.56</td>
<td>1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>33</td>
<td>Did you have confidence in the decisions made about your condition or treatment?</td>
<td>2</td>
<td>-0.78</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>34</td>
<td>How much information about your condition or treatment was given to you?</td>
<td>2</td>
<td>1.48</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>35</td>
<td>Did you find someone on the hospital staff to talk to about your worries and fears?</td>
<td>2</td>
<td>-0.77</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>36</td>
<td>Do you feel you got enough emotional support from hospital staff during your stay?</td>
<td>2</td>
<td>-0.75</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>37</td>
<td>Were you given enough privacy when discussing your condition or treatment?</td>
<td>2</td>
<td>1.58</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>38</td>
<td>Were you given enough privacy when being examined or treated?</td>
<td>2</td>
<td>-0.65</td>
<td>1.01</td>
<td>1.02</td>
</tr>
</tbody>
</table>

*aThreshold denotes the number of categories on each item, for example, 2 for three categories and 1 for two categories.

bDifficulty represents item difficulty in a unit of logit (=log odds).

### Task 1: Investigating Computer Adaptive Testing Efficiency and Accuracy

Table 2 indicates that the CAT relates the true scores ($r=.97$ in column 2) and the AAI ($r=.97$ in columns 3 and 4) to a high association, indicating that the CAT earns an equivalent accuracy compared to the AAI and a higher accuracy than the RSM (in column 5 of the estimation section). The summation scores have a higher correlation ($r=.98$) to the within counterparts (eg, summation RSM scores vs estimated RSM logit scores) and a slightly lower correlation ($r=.92-.97$) to the between counterparts (eg, summation RSM scores vs estimated AAI or CAT logit scores), implying that the raw summation scores have a high correlation ($r=.98$) with the estimated logit scores shown in the last 4 columns of Table 2. The bottom row of Table 2 shows that the CAT earns the shortest item length, indicating the CAT has advantages in efficiency over AAI and RSM.
Table 2. Comparisons of efficiency and accuracy among the AAI, RSM, and CAT.

<table>
<thead>
<tr>
<th></th>
<th>Estimated logit scores</th>
<th>Summation scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True score</td>
<td>CAT</td>
</tr>
<tr>
<td>True score</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>Estimation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.97</td>
<td>[0.98]</td>
</tr>
<tr>
<td>AAI</td>
<td>0.96</td>
<td>[0.98]</td>
</tr>
<tr>
<td>RSM</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Summation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSM</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>AAI</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>Mean</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td>Item length</td>
<td></td>
<td>40.50</td>
</tr>
</tbody>
</table>

aItem length denotes those items, excluding the 9 automatic-selection-to-different-path items.

**Task 2: Online Computer Adaptive Testing Assessment**

By scanning the QR code (see Figure 2), the CAT icon appears on the patient’s mobile phone. The mobile CAT survey procedure was demonstrated item-by-item in action (see Figure 3). Person fit (ie, infit and outfit MNSQ) statistics showed the respondent behaviors. Person theta is the provisional ability estimated by the CAT module.

The standard error in Figure 3 was generated by the following formula (see Multimedia Appendix 5): $1/\sqrt{\sum \text{variance}(i)}$, where $i$ refers to the CAT finished items responded to by a person [30]. In addition, the residual (resi) in Figure 3 was the average of the last five change differences between the pre-and post-estimated abilities on each CAT step. CAT will stop if the residual value is less than 0.05. “Corr” refers to the correlation coefficient between the CAT estimated measures and its step series numbers using the last five estimated theta (=person measure) values. The flatter the theta trend, the higher the probability that the person measure is convergent with a final estimation.
Figure 2. A snapshot of a QR code and the CAT item.

Figure 3. The process of CAT estimated scores.

Did doctors talk in front of you as if you weren’t there?

Yes, often

Yes, sometimes
Discussion

Principal Findings

We verified that computer adaptive testing can be (1) efficient with questionnaire item reduction and (2) used with QR codes on mobile phones.

The CAT item pool was designed using Items 15-38 of the UK NHS 70-item questionnaire. We found that CAT can be more efficient for answering questions than both AAI and RSM without compromising its measurement accuracy, which is consistent with previous studies [6-9]. Our online CAT inpatient survey for gathering satisfaction responses from patients was accessed by scanning a QR code on a mobile phone, which has never been demonstrated previously.

Many studies have discussed patient perceptions about hospitals and the benefit of listening to other patient experiences when choosing a hospital [31,32]. There has been a rapid increase in the number of websites that allow patients to rate their hospitals [33,34]. Almost all health care providers have been explicitly required to conduct surveys of their patients’ health care experiences. However, those surveys often use an individual item-by-item approach to disclose patient views on hospital service quality, which does not provide hospital staff with information to make further improvements without considering the overall hospital performance [13].

Implications and Future Considerations

We demonstrated that an NHS inpatient experience questionnaire with shortened items can be used with an IRT-based CAT technique without compromising its measurement accuracy. Using a CAT approach with such complex question structure jointly with item pools and conditional path skip items is rare. Our online CAT module used by scanning a QR code on a mobile phone can be extended to many dimensions simultaneously in a survey. For example, the Clinical Dementia Rating scale [35] used in patients with dementia consists of six domains. We could design a module using CAT through several procedures in a common questionnaire in the future.

Strengths

Hospital staff must consider both the efficiency and utility of assessment for the selection of the CAT items [36]. The traditional survey collects all feedback from patients through particular sets of questions to assess what causes patient difficulty or dissatisfaction. The assessment results help hospital managers determine where improvements can be made [36]. We can use the Rasch simulation technique to overcome the problem in questionnaires of unanswered items (ie, which do not provide hospital staff with information to make further improvement). This Rasch simulation technique [21] can be used to fill in the expected responses to those unanswered CAT items according to the final person theta (ability) and the specified item difficulties. Thus, the CAT can provide efficient assessments and the full information needed to make improvements.

Furthermore, the person outfit mean square in CAT is also saved in our database (see Multimedia Appendix 3). An outfit mean square of 2.0 or greater for a patient indicates a possibly aberrant response pattern [37], such as cheating, careless responding, lucky guessing, creative responding, or random responding [38], which makes it hard to reveal valuable information using the traditional survey method.

Limitations

Six limitations of this study are addressed. First, the study was based on the assumption of unidimensionality across those 24 CAT items. Although several articles have supported the notion that the UK NHS 70-item questionnaire can construct a one-dimension domain [13,31], those items cannot be generalized to the 24 CAT items used in different countries or by different groups. Future studies should further verify those 24 items to make the CAT module valid and feasible in health care practice.

Second, the first CAT item was selected from a randomized item pool. The CAT selection rule for the first item can be redesigned referring to the previously completed items and inferring a provisional theta (ie, person measure) to select the first item with the maximum information (ie, variance) in the item pool so that the questionnaire length could be shorter (see Multimedia Appendix 6).

Third, only one CAT module was designed in the NHS inpatient questionnaire due to the conditional selection path skip items that existed in non-CAT items. Future studies are recommended to overcome this barrier and to design a CAT-by-CAT approach in the questionnaire so as to reduce more item length in a questionnaire.

Fourth, we have not discussed the issue of participation options using traditional postal mail or email. Because not all patients possess a mobile phone, specifically a smartphone, and 3G/4G WiFi communication, all options (mail or email) must be offered to patients when invited to participate in the survey. Readers have found the email option useful to answer questions either by connecting to the Web, or by scanning a QR code on a mobile phone (eg, Figure 2) if applying the CAT demonstrated in this study. Future studies are needed to further explore and improve the processes of the CAT survey.

Fifth, we conducted a simulation of 1000 patients’ true scores followed by a standard normal distribution. This might contradict the general experience of satisfaction surveys with ceiling effects that impede standard normal distribution. Future studies are needed to sample from a negatively skewed population to further verify whether the CAT can be efficient on item reduction over AAI and RSM.

Sixth, the unidimensionality of the 24 CAT items may be questioned given the different realms of hospital ward, doctors, nurses, patient care, and treatment. It might be implausible to assume that these 24 items are unidimensional. Future studies are required to further investigate the issue.

Conclusion

With advances in technology, we can now offer patients alternative ways via mobile phones to gather their feedback on.
hospitalization satisfaction. The online CAT can reduce the number of survey items for patients to respond to, as well as be accessed via mobile phone using a QR code.

**Authors' Contributions**

All authors have read and approved the final manuscript. TW developed the study concept and design. TW and WP analyzed and interpreted the data. TW drafted the manuscript, and all authors have provided critical revisions for important intellectual content. The study was supervised by TW.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**

A format of skip items in a questionnaire.

[PDF File (Adobe PDF File), 224KB - medinform_v4i1e8_app1.pdf]

**Multimedia Appendix 2**

The UK NHS 70-item questionnaire.

[PDF File (Adobe PDF File), 269KB - medinform_v4i1e8_app2.pdf]

**Multimedia Appendix 3**

The study item dataset.

[XLS File (Microsoft Excel File), 14MB - medinform_v4i1e8_app3.xls]

**Multimedia Appendix 4**

Excel VBA module used for CAT simulation.

[XLS File (Microsoft Excel File), 7MB - medinform_v4i1e8_app4.xls]

**Multimedia Appendix 5**

Excel module for simulation on CAT, AAI, and RSM saving to spreadsheet NAT.

[TXT File, 32KB - medinform_v4i1e8_app5.txt]

**Multimedia Appendix 6**

Introduction to CAT.

[PDF File (Adobe PDF File), 405KB - medinform_v4i1e8_app6.pdf]

**References**


**Abbreviations**

- AAI: answering all items
- CAT: computer adaptive testing
- HCAHPS: Hospital Consumer Assessment of Healthcare Providers and Services
- IRT: item response theory
- MNSQ: mean square
- NHS: National Health Service
- PCM: partial credit model
- QR: quick response
- RSM: randomized selection method
- SE: standard error
- SEM: standard error measurement
- VBA: Visual Basic for Applications

©Tsair-Wei Chien, Weir-Sen Lin. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 02.03.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.
Computerized Automated Quantification of Subcutaneous and Visceral Adipose Tissue From Computed Tomography Scans: Development and Validation Study

Abstract

Background: Computed tomography (CT) is often viewed as one of the most accurate methods for measuring visceral adipose tissue (VAT). However, measuring VAT and subcutaneous adipose tissue (SAT) from CT is a time-consuming and tedious process. Thus, evaluating patients’ obesity levels during clinical trials using CT scans is both cumbersome and limiting.

Objective: To describe an image-processing-based and automated method for measuring adipose tissue in the entire abdominal region.

Methods: The method detects SAT and VAT levels using a separation mask based on muscles of the human body. The separation mask is the region that minimizes the unnecessary space between a closed path and muscle area. In addition, a correction mask, based on bones, corrects the error in VAT.

Results: To validate the method, the volume of total adipose tissue (TAT), SAT, and VAT were measured for a total of 100 CTs using the automated method, and the results compared with those from manual measurements obtained by 2 experts. Dice’s similarity coefficients (DSCs) between the first manual measurement and the automated result for TAT, SAT, and VAT are 0.99, 0.98, and 0.97, respectively. The DSCs between the second manual measurement and the automated result for TAT, SAT, and VAT are 0.98, 0.98, and 0.97, respectively. Moreover, intraclass correlation coefficients (ICCs) between the automated method and the results of the manual measurements indicate high reliability as the ICCs for the items are all .99 (P<.001).

Conclusions: The results described in this paper confirm the accuracy and reliability of the proposed method. The method is expected to be both convenient and useful in the clinical evaluation and study of obesity in patients who require SAT and VAT measurements.

(JMIR Med Inform 2016;4(1):e2) doi:10.2196/medinform.4923

KEYWORDS

obesity; visceral adipose tissue; subcutaneous adipose tissue; computed tomography; computer-assisted image analysis

Introduction

Obesity refers to the over-accumulation of adipose tissue (AT) in the body. Obesity can be caused by genetic and fat metabolism abnormalities, hypothyroidism, excessive nutritional intake, lack of exercise, and stress, and it is known to be a key factor in chronic diseases [1]. In recent studies, body fat distribution has been shown to pose a greater health risk than overall body fat, and among the different types of obesity based on specific categories of body fat distribution, abdominal obesity has been reported to pose the greatest risk [2]. Abdominal obesity increases the prevalence rate of metabolic syndrome accompanied by coronary artery diseases such as hypertension, diabetes, hyperlipidemia, and arteriosclerosis [3,4]. Because visceral adipose tissue (VAT), rather than subcutaneous adipose tissue (SAT), is recognized as the contributing factor in body insulin resistance, visceral abdominal obesity is viewed as the...
more clinically important type of abdominal obesity [5]. Furthermore, by causing physical pressure, the accumulation of heavy VAT can interrupt blood flow to abdominal organs and decrease organ function (eg, liver). As such, VAT can be even more deleterious than SAT [6]. The accurate evaluation and prevention of both SAT and VAT with quantitative fat measurements are thus important.

In recent clinical trials, diverse methods have been applied to assess obesity. For example, body mass index (BMI) can be used to evaluate obesity easily by height and weight measurements; however, VAT cannot be measured by this method. On the other hand, methods such as bioelectrical impedance analysis (BIA), magnetic resonance image (MRI), and computed tomography (CT) can acquire VAT measurements. BIA is a technique by which the percentage of body fat and the area of VAT are estimated by sending a weak-to-high-frequency current through the body and measuring the bioelectrical impedance [7]. Recently, BIA has been widely used in diagnosing obesity owing to the simplicity of its measurement; many reports support its high-degree of accuracy in body fat mass measurements [8]. However, there are only a few reports on whether it can satisfactorily reflect the actual amount of VAT. On the other hand, CT can clearly distinguish AT from other tissues, and AT can be measured directly from the cross-sectional images of the tissue. In addition, CT has the advantage that SAT and VAT can also be directly measured [9]. Because the measurement method is cumbersome and poses a risk of radiation exposure, it is not in general use. However, the additional radiation exposure involved in measuring AT can be eliminated by using CT images obtained from health screening or other procedures, and by using low-dose CT, which is currently being used in the clinic. Finally, MRI, which is similar to CT, can be used to generate cross-sectional images and can measure tissue and VAT directly. Despite the advantage of no radiation exposure [10], MRI has clinical limitations because of the cumbersome measuring method, high-cost, and lengthy imaging times.

Among the various obesity evaluation methods, CT is considered the gold standard in clinical trials due to its high accuracy. However, AT has to be measured directly from the cross-sectional images, and when measuring VAT, the boundaries between SAT and VAT must be clearly defined. In other words, a substantial amount of time and effort have to be expended to analyze a single CT image. In abdominal obesity-related studies, such problems can act as limiting factors in analyzing large amounts of data or when analyzing a wide range of AT in the abdomen. Efforts have been made to minimize the time required to make measurements by calculating the level of abdominal obesity from a single CT image at the umbilical level (L4-5 vertebrae), which represents the entire abdominal fat area [11].

In this study, we attempt to solve these limiting factors by using an automated, computer image processing technique. By automating the image processing, the entire abdominal region can be measured in a relatively short time period. In addition, it eliminates subjectivity, enabling objective, quantitative, and reliable measurements to be made. An increasing number of studies have been conducted on the computerized, automated measurement of AT. For example, using a single CT image at the umbilical level, Bandekar et al proposed the use of the active shape model and fuzzy affinity-based automatic fat analysis to distinguish between SAT and VAT [12]. In the study, they compared the results with those obtained manually by experts, and they calculated and evaluated both accuracy and sensitivity. They determined the degree of accuracy for SAT and VAT to be 98.29% (SD 0.62%) and 97.66% (SD 0.98%), respectively [12]. Zhao et al also reported automated separation of SAT and VAT using pixel information obtained from radial movement at increments of 3 degrees from the center of the body from a single CT image at the umbilical level [13]. In that study, 9 subjects were tested and the differences between the automated and manual measurements in terms of SAT and VAT were 0.65% and 1.54%, respectively [13]. In another example, Kullberg et al used a histogram based on MRI images at the umbilical level range and constructed AT and SAT masks to isolate each AT [14]. An algorithm was applied to a total of 17 data values, using the manual measurements as the reference, and obtained true positive measurements for SAT and VAT, with high accuracy values of 96% (SD 2.3%) and 90% (SD 6.5%), respectively [14]. Although the studies mentioned above showed a high degree of accuracy, the verification data was minimal, and the allowed range of measurement was limited to a single image or restricted to the umbilical level only. In order to overcome these limitations, we propose a method for the automatic separation and measurement of SAT and VAT in the entire abdominal region using CT. By comparing a large amount of test data with manually measured results, the technical and clinical utility of the proposed method was verified and evaluated.

**Methods**

**Study Dataset and Development Environment**

This study was approved by the Institutional Review Board (IRB) of the National Cancer Center of Korea with a waiver of the requirement for patients’ informed consent (1210160-3).

In this study, abdominal CTs on 100 patients from the National Cancer Center of Korea were analyzed. Images were obtained using LightSpeed VCT (GE Healthcare) and Brilliance 64 (Phillips). For each CT scan, an image size of 512 × 512 pixels and slice thicknesses of 2.5 to 5.0 mm were used. We measured the abdominal area, which included the region from the diaphragm to coccyx and excluded the pelvic cavity. Here, Microsoft Virtual Studio (Ver. 2005, Microsoft, Redmond, US) was used for the development of algorithms and software, and ITK (Ver. 3.14.0, Kitware, US) and VTK (Ver. 5.10.0, Kitware, US) were used as libraries. The SPSS package (Ver. 13, SPSS Inc., US) was used for statistical analysis.

**Manual Measurement**

The manual measurements reported in this paper were taken directly by 2 experts using in-house developed software. The in-house software depicts the AT area using a brush to color the area directly on the CT image; SAT and VAT are distinguished by applying different colors to each. For the convenience of the user, only areas within the AT attenuation range of -30 to -190 Hounsfield units (HU) in CT and...
multi-planar reconstruction were considered [11,15]. The manual measurement results were used as the gold standard data in comparisons with the automated measurement results.

**Automatic Measurement**

The automated method of body fat proposed measurement described in this paper consists of the following three tasks: (1) pre-processing, (2) fat detection, and (3) post-processing. The complete flowchart of the algorithm is shown in Figure 1.

**Preprocessing**

CT images include various objects such as body, air, bed, and sheets. The efficiency of the algorithm was enhanced to perform the analysis only within the body region. The improved algorithm was applied to the pre-processing stage for the body-region detection. Unnecessary regions of the CT image were removed to prevent errors in detecting the body region. The body-region detection was performed using threshold and labeling techniques [16]. The air’s HU in CT images has attenuation lower than -1000 HU [17]. Knowing this, we eliminated the air region by setting thresholds and detected other regions by labeling. Subsequently, the body area was acquired by all the other labels, except the label with the widest area.

**Fat Detection**

AT in the abdominal area is sorted in SAT and VAT according to the location of AT and abdominal muscles; the interior of muscle is classified as SAT and the exterior as VAT. Thus, a separate mask was created based on the location of the abdominal muscles.

Creating a separation mask consists of three steps: (1) non-fat area segmentation; (2) closed path acquisition; (3) and correction of the closed path (Figures 2-4). Non-fat area detection is a process for finding the location of the muscle as the base of a separation mask and is performed by eliminating the abdominal fat region through thresholds. First, total adipose tissue (TAT) is obtained by setting a threshold at -30 to -190 HU, the attenuation range of fat in CT images, and the non-fat area is...
extracted by removing the TAT area from the body area [18]. After extraction, the skin area included in the area using opening, a morphology-based technique, is deleted. The result of this process shows that bones and organs inside the abdominal cavity as well as muscles are detected in the area (Figure 2). Muscle segmentation is omitted because bones and organs inside muscle don’t affect a closed path, which is detected on the basis of the outermost muscle coordinates. In addition, the efficiency of the algorithm is increased by skipping the muscle segmentation process. Obtaining a closed path is the process for blocking the parts connected between SAT and VAT completely, such as the ones shown in Figure 2. The Convex Hull algorithm was used to detect the shortest closed path [19]. This algorithm determines the convex polygon of the minimum area included when a set of points or shapes are given. The Convex Hull results show the complete separation of organ areas without connecting the parts between SAT and VAT. However, errors can be generated as a closed path cannot be perfectly attached to muscle area when detecting a closed path (Figure 3). Because the errors can affect the quantitative results, each coordinate is corrected so that a closed path can be closer to the muscle area. Correction of errors is performed on all the coordinates that make up a closed path and the procedure is shown in Textbox 1.

Textbox 1. Procedure for the correction of errors on the coordinates that make up a closed path.

1. Examine whether the closed path is in contact with the organ region relative to the y axis coordinates of the relevant closed path.
2. If it is not in contact, the relevant coordinate is deleted and moved to the location where it is in contact with the organ region.
3. After completing the examination of all closed path coordinates, randomly identify 1 coordinate as the starting point.
4. With the starting point as the standard, set the coordinate located at the shortest distance as the arrival point. After connecting it with a line, set the arrival point as the new starting point.
5. Repeat Stage 4 until it cannot be performed further.

The lines drawn through this method of coordinate correction become the new closed path of the organ region and produce the separation mask by filling in the interior of the closed path. The separation mask’s outline can be verified in Figure 4 and the entire coordinate correction process can be observed in Figure 5.

Figure 2. Result of the threshold application to the visceral region.

Figure 3. Result of the application of the Convex Hull algorithm and threshold.

The produced separation mask sorts and detects for each region of SAT and VAT. SAT regions can be extracted by removing the separation mask from the TAT region detected by producing the separation mask. VAT regions can be detected through an “AND” operation between TAT and the mask regions.
Post-Processing

The fat component inside bones is detected as body fat in the process of detecting TAT, and it is included in the separation mask as bones are surrounded by muscles. Consequently, it is incorrectly detected as VAT. The spine, ribs, and pelvis surrounding the right and left abdominal cavity from the center of the back, and AT need be erased. In order to eliminate the false detection as VAT, a correction mask based on bones was included. Bones have a pixel range above 1000 HU in CT [17], and based on that, we set a seed point in the region estimated as bone and detected them using a 3D region growing algorithm [20]. Then, a correction mask was made by obtaining a closed path and applying a closed path-correction method to the extracted bone region in the same way as the separation mask. We attempted to apply different correction methods for each region after dividing the body into 2 regions based on bones. The regions are divided considering whether the right or left side of the abdominal cavity is surrounded by bones. The region that the ribs and pelvis are located in is surrounded on the back, right, and left sides of the abdominal cavity by bones. Therefore, it is necessary consider all of the sides of the abdominal cavity (Area 1). However, only the region of the back side that contains the spine needs to be imaged (Area 2). Considering these differences, the region detected as VAT outside a separation mask (back side, right and left side), as well as inside a separation mask for Area 1 were eliminated. The region recognized by the VAT in Area 2 in the back, outside of the separation mask, as well as inside the separation mask was removed. Any false-detected VAT was corrected using Area 1 and Area 2.

Results

In this study, we detected and measured the volume of TAT, SAT, and VAT for a total of 100 CT data using our automated method (Figure 6). As well, the automated measurement results ($M_{AUT}$) were compared with the manual measurement results obtained by 2 experts ($M_{M1}$, $M_{M2}$). A comparative analysis was carried out between them in order to verify the technical accuracy and clinical reliability of the proposed method. The accuracy of the automated measurement method was evaluated through four kinds of conditional probability including sensitivity, specificity, accuracy, and Dice’s similarity coefficient (DSC). The automated measurement results and the manual measurement results were also compared through one-way analysis of variance (ANOVA) tests and Bland-Altman plots (see Multimedia Appendix 1), performing regression analysis and scatter plots (see Multimedia Appendix 2) to investigate the correlation between them. Moreover, the intraclass correlation coefficients (ICCs) between the results of both measurement methods were determined, and the reliability between the results examined. For the conditional probability test, true positive, false positive, true negative, and false negative
were obtained by calculating, pixel-by-pixel, the position of VAT, SAT, and TAT detected by the automated and manual methods. The test results are shown in Table 1. The test results of M1 indicate high precision, as the accuracy of M1 for TAT, SAT, and VAT is 99.69%, 99.79%, and 99.79%, respectively. The DSC values for TAT, SAT, and VAT are 0.99, 0.98, and 0.97, respectively. Similar results were obtained for M2, as the accuracy and DSC value for TAT is 99.62% and 0.98, 99.77% and 0.98 for SAT, and 99.74% and 0.97 for VAT.

**Table 1.** The conditional probability test between the automated and manual measurements.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Accuracy, %</th>
<th>Dice similarity coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automatic measurement (M AUT) and manual measurement (M M1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAT</td>
<td>97.45</td>
<td>99.96</td>
<td>99.69</td>
<td>0.99</td>
</tr>
<tr>
<td>SAT</td>
<td>97.24</td>
<td>99.98</td>
<td>99.79</td>
<td>0.98</td>
</tr>
<tr>
<td>VAT</td>
<td>97.54</td>
<td>99.88</td>
<td>99.79</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>Automatic measurement (M AUT) and manual measurement (M M2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAT</td>
<td>97.39</td>
<td>99.89</td>
<td>99.62</td>
<td>0.98</td>
</tr>
<tr>
<td>SAT</td>
<td>96.96</td>
<td>99.98</td>
<td>99.77</td>
<td>0.98</td>
</tr>
<tr>
<td>VAT</td>
<td>97.87</td>
<td>99.81</td>
<td>99.74</td>
<td>0.97</td>
</tr>
</tbody>
</table>

The mean volume of TAT, SAT, and VAT measured by the automated method was 7913.79 mL (SD 2852.62 mL), 4620.38 mL (SD 1735.76 mL), and 3293.41 mL (SD 1497.11 mL), respectively. The mean volume for the same items measured by the manual method was 8021.56 mL (SD 2877.91 mL), 4750.01 mL (SD 1801.47 mL), and 3271.54 mL (SD 1469.16 mL) for M1, and 7972.33 mL (SD 2899.43 mL), 4757.41 mL (SD 1822.06 mL), and 3214.91 mL (SD 1473.27 mL) for M2 (Table 2). The correlations are significant between the volumes of M AUT and M M1 for TAT (r=.999, P<.001), SAT (r=.999, P<.001), and VAT (r=.999, P<.001) (Multimedia Appendices 2A-C). The correlations between the volumes of M AUT and M M2 are also significant for TAT (r=.999, P<.001), SAT (r=.999, P<.001), and VAT (r=.999, P<.001) (Multimedia Appendices 2D-F).

**Table 2.** Comparison and verification between the results of the automated and manual measurements.

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean volume(^a) (SD)</th>
<th>F</th>
<th>P value(^b)</th>
<th>ICC</th>
<th>P value(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAT</td>
<td></td>
<td>.035</td>
<td>.965</td>
<td>.99</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>(M_{AUT})</td>
<td>7913.79 (2852.62)</td>
<td>(M_{M1})</td>
<td>8021.56 (2877.91)</td>
<td>(M_{M2})</td>
<td>7972.33 (2899.43)</td>
</tr>
<tr>
<td>(M_{AUT})</td>
<td>4620.38 (1735.76)</td>
<td>(M_{M1})</td>
<td>4750.01 (1801.47)</td>
<td>(M_{M2})</td>
<td>4757.41 (1822.06)</td>
</tr>
<tr>
<td>SAT</td>
<td>(M_{AUT})</td>
<td>.186</td>
<td>.830</td>
<td>.99</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>(M_{M1})</td>
<td>3293.41 (1497.11)</td>
<td>(M_{M2})</td>
<td>3271.54 (1469.16)</td>
<td>(M_{M2})</td>
<td>3214.91 (1473.27)</td>
</tr>
</tbody>
</table>

\(^a\) Mean volumes and SD measured in milliliter.
\(^b\) P value for ANOVA test.
\(^c\) P value for ICC.

One-way ANOVA test results for the volumes of \(M_{AUT}\), \(M_{M1}\), and \(M_{M2}\) revealed that no significant differences were found for TAT (F=.035, P=.965), SAT (F=.186, P=.830), and VAT (F=.075, P=.928; Table 2). Bland-Altman plots for the same items showed a good comparability as most volumes were within 1.96 standard deviations from the average position of the
respective volume differences (Multimedia Appendices 1A-F). ICC values for the volumes of $M_{AUT}$, $M_{M1}$, and $M_{M2}$ indicate high reliability for all the measuring items with .99 (CI 0.99-0.99, $P<.001$) for TAT, .99 (CI 0.97-0.99, $P<.001$) for SAT, and .99 (CI 0.98-0.99, $P<.001$) for VAT.

The elapsed time required to measure constant volume with both the automated and manual methods was compared to evaluate the usefulness of the proposed method. To do this, 20 data sets were randomly selected out of the 100 total, each containing a slice thickness of 5 mm. The elapsed time required to determine TAT, SAT, and VAT, within the range of the umbilical level (24 12 cm slices) by the automated and the manual method was measured. We found that the mean elapsed time required for manual measurements performed by 2 experts was 718.5 seconds (SD 72.7 seconds) for $M_{M1}$ and 815.5 seconds (SD 65.8 seconds) for $M_{M2}$, whereas the automated method ($M_{AUT}$) required only 3.62 seconds (SD 0.1 seconds) (Figure 7).

![Figure 6. Automated segmentation and measurement results for SAT (green) and VAT (red).](image)

![Figure 7. Comparison of the elapsed time required for measuring body fat within the range of the umbilical level (12 cm).](image)

**Discussion**

**Principal Findings**

The previous AT measurement methods with CT were very tedious and time consuming. As such, many studies have attempted to find out the association between the indicators of obesity such as BMI, waist to hip ratio (WHR), and fat measurements with the typical CT used in clinic [21,22]. Although one CT has some association with diverse indicators of obesity, there’s a risk of error caused by measurements and evaluations of a 2-dimensional (2D) cross section of a 3-dimensional (3D) body. In addition, indicators such as BMI and WHR are not quantitative. Therefore, we proposed and
verified a new method of separating and measuring SAT and VAT from the entire abdominal CT automatically by adopting an image processing technique. The test results of the proposed method demonstrate a high level of accuracy (99%) for TAT, SAT, and VAT compared to the results made manually (M1 and M2).

The results of our proposed method show a higher degree of accuracy compared with other reported studies (Table 3). Bandekar et al [12] applied the active shape model automatic fat methods based on fuzzy affinity and reported accuracies of 98.29% (SD 0.62%) and 97.66% (SD 0.98%) for SAT and VAT, respectively. Using a method of obtaining pixel information by a 3 degree radial movement from the body center to understand SAT and VAT on CT images, Zhao et al [13] reported accuracies of 99.35% for SAT and 98.46% for VAT. In another example, Kullberg et al [14] measured accuracies of 96% (SD 2.3%) for SAT and 90% (SD 6.5%) and VAT using a histogram based on MRI images in the umbilical level range and constructed AT and SAT masks to isolate each AT.

Table 3. Comparison of the automated measurement of AT proposed in this paper with previously published studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Modality</th>
<th>Range</th>
<th>Number of data sets</th>
<th>Accuracy of SAT, % (SD)</th>
<th>Accuracy of VAT, % (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandekar et al [12]</td>
<td>CT</td>
<td>1 slice at umbilical level</td>
<td>40</td>
<td>98.29 (0.62)</td>
<td>97.66 (0.98)</td>
</tr>
<tr>
<td>Zhao et al [13]</td>
<td>CT</td>
<td>1 slice at umbilical level</td>
<td>9</td>
<td>99.35</td>
<td>98.46</td>
</tr>
<tr>
<td>Kullberg et al [14]</td>
<td>MRI</td>
<td>Volume of umbilical level</td>
<td>17</td>
<td>96 (2.3)</td>
<td>90 (6.5)</td>
</tr>
<tr>
<td>Proposed method</td>
<td>CT</td>
<td>Volume of entire abdominal cavity</td>
<td>100</td>
<td>99.78 (0.18)</td>
<td>99.76 (0.16)</td>
</tr>
</tbody>
</table>

With respect to measurement range, our proposed method can measure the entire abdominal cavity rather than the existing methods that are restricted to 2-dimensional levels or only the umbilical region. While the way of separating SAT and VAT using templates or masks is similar to other studies, in our proposed method, the separation mask is generated by minimizing the number of algorithms based on pixel value, approaching the anatomical shape. Furthermore, the analysis is a contributing factor for raising the level of accuracy.

The ANOVA test and ICC results demonstrate the clinical reliability of the proposed method (Table 2). The ANOVA test indicates that no noticeable differences were observed between the automated measurements and the manual measurements made by the 2 experts for TAT, SAT, and VAT. The ICC results show a very high level of reliability for TAT, SAT, and VAT (.99). However, we did find that the automated measurements had slightly smaller values than the manual measurement for SAT, while for VAT, the automated measures had the tendency to be slightly larger. A possible cause for the gap in SAT segmentation is because intermuscular fat or intramuscular fat near subcutaneous fat is included as part of subcutaneous fat, and it is accumulated during the manual measurement. The difference in VAT segmentation might be caused by the detection of some adipose tissue inside the bone as VAT during the correction process of the automated measurement. Although their impact on the overall segmentation results was negligible since the rate of error was low, an even higher degree of accuracy can be anticipated as the algorithms are improved in the future. The time-shortening effect of the automated method shows the time reduction to be about 200 fold compared to the manual method for a constant range (Figure 7). Based on this result, we can estimate that the time required to manually measure an abdominal CT consisting of 150 images with a 3 mm thickness to be approximately 80 min while the same can be done using the automated method in approximately 20 seconds. This supports the usefulness and power of the automated method. Although the proposed method uses CT for measurement, the structure of the algorithm is based on anatomical shape. Therefore, the potential exists to apply it to other imaging methods such as MRI.

Conclusions
Measuring AT by the proposed CT method allows for a more quantitative and objective measurement result in a time efficient manner. As well, the simple measurements of VAT enabled by the new method will be very useful in evaluating visceral abdominal obesity. Furthermore, we expect that the proposed method will be more convenient and useful in clinical evaluations and studies on patients’ abdominal obesity levels.

Acknowledgments
This work was supported by a research grant from the National Cancer Center, Korea (grant 1610050-1).

Conflicts of Interest
None declared.
References


**Abbreviations**

- **ANOVA**: analysis of variance
- **AT**: adipose tissue
- **BIA**: bioelectrical impedance analysis
- **CT**: computed tomography
- **DSC**: Dice’s similarity coefficient
- **HU**: Hounsfield unit
- **ICC**: intraclass correlation coefficient
- **MRI**: magnetic resonance image
- **SAT**: subcutaneous adipose tissue
- **TAT**: total adipose tissue
- **VAT**: visceral adipose tissue
- **WHR**: waist-to-hip ratio

©Young Jae Kim, Ji Won Park, Jong Wan Kim, Chan-Soo Park, John Paul S Gonzalez, Seung Hyun Lee, Kwang Gi Kim, Jae Hwan Oh. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 04.02.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.