Contents

Original Papers

Types and Frequency of Infusion Pump Alarms and Infusion-Interruption to Infusion-Recovery Times for Critical Short Half-Life Infusions: Retrospective Data Analysis (e14123)
James Waterson, Arkadiusz Bedner. .......................................................... 2

Postinjury Complications: Retrospective Study of Causative Factors (e14819)
Elizabeth Warnack, Hersch Pachter, Beatrix Choi, Charles DiMaggio, Spiros Frangos, Michael Klein, Marko Bukur. ............................................. 35

Evaluating the Feasibility of a Software Prototype Supporting the Management of Multimorbid Seniors: Mixed Methods Study in General Practices (e12695)
Christine Kersting, Birgitta Weltermann. .................................................. 45

Understanding the Situated Roles of Electronic Medical Record Systems to Enable Redesign: Mixed Methods Study (e13812)
Samar Helou, Victoria Abou-Khalil, Goshiro Yamamoto, Eiji Kondoh, Hiroshi Tamura, Shusuke Hiragi, Osamu Sugiya, Kazuya Okamoto, Masayuki Nambu, Tomohiro Kuroda. ................................................................. 55

Factors Associated With Electronic Health Record Usage Among Primary Care Physicians After Hours: Retrospective Cohort Study (e13779)
Selasi Attipoe, Yungui Huang, Sharon Schweikhart, Steve Rust, Jeffrey Hoffman, Simon Lin. .......................................................... 73

Health Care Professionals' Clinical Perspectives and Acceptance of a Blood Glucose Meter and Mobile App Featuring a Dynamic Color Range Indicator and Blood Sugar Mentor: Online Evaluation in Seven Countries (e13847)
Mike Grady, Usha Venugopal, Katia Robert, Graham Hurrell, Oliver Schnell. .......................................................... 79

Tailoring of a Smartphone Smoking Cessation App (Kick.it) for Serious Mental Illness Populations: Qualitative Study (e14023)
Paucline Klein, Sharon Lawn, George Tsourtos, Joep van Agteren. .......................................................... 91

Review

Social Cognitive Theories and Electronic Health Design: Scoping Review (e11544)
Patricia Grace-Farfaglia. .......................................................... 14
Types and Frequency of Infusion Pump Alarms and Infusion-Interruption to Infusion-Recovery Times for Critical Short Half-Life Infusions: Retrospective Data Analysis

James Waterson1*, BA, MMedEd; Arkadiusz Bedner1*, BSc, MSc
Medication Management Solutions, Becton Dickinson Limited, Eysins, Switzerland

*all authors contributed equally

Corresponding Author:
James Waterson, BA, MMedEd
Medication Management Solutions
Becton Dickinson Limited
BD Switzerland Sàrl
Eysins, 1262 Switzerland
Phone: 41 215563128
Email: jessjimw@yahoo.com

Abstract

Background: Alarm fatigue commonly leads to a reduced response to alarms. Appropriate and timely response to intravenous pump alarms is crucial to infusion continuity. The difficulty of filtering out critical short half-life infusion alarms from nonurgent alarms is a key challenge for risk management for clinicians. Critical care areas provide ample opportunities for intravenous medication error with the frequent administration of high-alert, critical short half-life infusions that require rigorous maintenance for continuity of delivery. Most serious medication errors in critical care occur during the execution of treatment, with performance-level failures outweighing rule-based or knowledge-based mistakes.

Objective: One objective of this study was to establish baseline data for the types and frequency of alarms that critical care clinicians are exposed to from a variety of infusion devices, including both large volume pumps and syringe drivers. Another objective was to identify the volume of these alarms that specifically relate to critical short half-life infusions and to evaluate user response times to alarms from infusion devices delivering these particular infusions.

Methods: The event logs of 1183 infusion pumps used in critical care environments and in general care areas within the European region were mined for a range of alarm states. The study then focused on a selection of infusion alarms from devices delivering critical short half-life infusions that would warrant rapid attention from clinicians in order to avoid potentially harmful prolonged infusion interruption. The reaction time of clinicians to infusion-interruption states and alarms for the selected critical short half-life infusions was then calculated.

Results: Initial analysis showed a mean average of 4.50 alarms per infusion in the general critical care pump population as opposed to the whole hospital rate of 1.39. In the pediatric intensive care unit (PICU) group, the alarms per infusion value was significantly above the mean average for all critical care areas, with 8.61 alarms per infusion. Infusion-interruption of critical short half-life infusions was found to be a significant problem in all areas of the general critical care pump population, with a significant number of downstream (ie, vein and access) occlusion events noted. While the mean and median response times to critical short half-life infusion interruptions were generally within the half-lives of the selected medications, there was a high prevalence of outliers in terms of reaction times for all the critical short half-life infusions studied.

Conclusions: This study gives an indication of what might be expected in critical care environments in terms of the volume of general infusion alarms and critical short half-life infusion alarms, as well as for clinician reaction times to critical short half-life infusion-interruption events. This study also identifies potentially problematic areas of the hospital for alarm fatigue and for particular issues of infusion and infusion-line management. Application of the proposed protocols can help create benchmarks for pump alarm management and clinician reaction times. These protocols can be applied to studies on the impact of alarm fatigue and for the evaluation of protocols, infusion-monitoring strategies, and infusion pump-based medication safety software aimed at reducing alarm fatigue and ensuring the maintenance of critical short half-life infusions. Given the frequency of infusion alarms seen in this study, the risk of alarm fatigue due to the white noise of pump alarms present in critical care, to which clinicians are...
Introduction

Background
Alarm fatigue is the presence of “frequent false alarms, leading to a reduced response to alarms” [1] and is caused by both false and nonactionable alerts. It can cause sleep disturbances, impaired healing, and intensive care unit (ICU) delirium for patients and contributes to staff burnout [2].

Infusion pumps are only one of numerous medical devices that are present at the bedside of critically ill adults, children, and neonates. Like other devices, there is a strong probability that their alarms may be ignored or at least downgraded in terms of the requirement for an immediate response by staff who are exposed to an excessive number of alarms, many of which are false, clinically insignificant, and essentially nonactionable [3]. Indeed, one study found that only 15% of cardiovascular alarms in an ICU setting were clinically relevant [4].

While alarm fatigue as a subject does embrace the alarms from all bedside medical devices, it is pertinent to focus on infusion pumps, as critical care patients commonly receive multiple infusions of sedatives, muscle relaxants, parental nutrition, fluid maintenance, antibiotics, and antivirals, as well as critical short half-life infusions. This means that the problem of white noise carries extra complications for these devices. Pump alarms first need to be separated out from other device alarms such as those from ventilators, monitors, and even low-grade technical alarms from beds or warming blankets. Alarms also need to be quickly rated for immediate response required as in the case of critical short half-life medications, where any infusion interruption can have potentially catastrophic hemodynamic consequences for the patient. These critical alarms requiring immediate response are only some of many other suitable nonimmediate response required alarms experienced by the clinician. For example, an alert for near end of infusion (NEOI) or end of infusion (EOI) requires an immediate response if the medication is a critical short half-life infusion but takes a lower priority if the alert is for a medication such as ganciclovir or amoxycillin completing a dose.

The issue at hand is that such filtering is extremely difficult with multiple infusions at the bedside and the problem is compounded when nurses are caring for more than one patient. This is made more complex by the common need to balance acuity in the nursing staff’s respective workload; giving each nurse an equal acuity load may require geographical convenience to be sacrificed. The increasingly common use of isolation rooms in critical care areas for infection control and protection of the patient also makes direct visualization of the patient and pumps increasingly difficult. Even beyond critical care, infusion therapy is one of the most widely practiced therapies in any health care organization, with alarms in the general hospital pump population accounting for 5% of lost infusion time per year [5].

A number of previous studies have focused on time and motion outcomes, device management, and equipment maintenance processes to reduce the number of alarms from pumps [5]. General reaction times have been recorded in one study where mean resolution times for 83% of alarms were 1 minute or less, but outliers of some 3% of alarms were not dealt with for more than 4 minutes [6].

Those studies are certainly of value and this study continues this work with a large-scale examination of alarm types. Added to this is the metric of reaction time for critical short half-life infusions and ready-made code for interrogation of the event logs of specific pumps.

Objectives
The overall objective of this study was to establish baseline data for volume of alarms and to note differences across the various environments that fall under critical care and within the data available to us from the following hospital units: pediatric intensive care unit (PICU), neonatal intensive care unit (NICU), general (adult) intensive care unit (ICU), and coronary (adult) intensive care unit (CICU). Each of these areas faces different challenges in terms of the complexity of infusion therapy, intravenous access devices used, patient characteristics, and staffing patterns.

In terms of critical short half-life infusion management and patient safety, the key measurements undertaken in the study included metrics to assist in an assessment of the ability of clinicians to differentiate between infusion-interruption alarms for these medications and other nonurgent alarms and to then act quickly to resolve infusion interruptions. The study’s objectives fell into two high-level processes:

1. Identifying the overall volume of pump alarms experienced versus the number of these alarms that apply to critical short...
half-life infusions is of value for the assessment of the background of white noise, against which clinicians are operating on a daily basis in the ICU environment.

2. Identifying clinician reaction time to critical short half-life infusion alarms and time to resolution of the issue.

The above information and the creation of protocols for interrogating pump event logs to gain this data was felt to be of value going forward for studies into strategies that could reduce spurious alarms, make prioritization of alarms easier, and improve reaction times to critical short half-life infusion interruption. In this respect, the value of measuring the total volume of alarms experienced by clinicians in critical care is a key parameter, as reducing this number overall would be expected to directly impact on the second measured parameter: reaction time to critical short half-life infusion interruption.

**Methods**

**Study Design**

This retrospective study was conducted on the anonymized event logs of 1183 infusion pumps—566 CareFusion, Becton, Dickinson and Company (BD) Alaris Plus and Guardrails Plus large volume pumps (LVPs) and 617 syringe pumps (SPs)—used in critical care environments within the European region over a period of 6482 days, between January 1, 2000, and September 29, 2017.

Pumps were of various ages and carried a range of dose error reduction software (DERS) solutions; in all cases, the pump's event log was bundled with near-miss medication error reporting data (ie, BD Alaris continuous quality improvement [CQI] software).

DERS software is generally comprised of two parts: (1) a drug library that contains the names of medications to be given by infusion, with minimum and maximum doses for each medication, and (2) a dataset that contains the drug library and has default configurations for the pump, including alarm settings such as maximum occlusion pressure, air-in-line threshold, and alarm volume. The features of the DERS used by the study pumps are given below.

The event logs were mined via SQL Server Management Studio v17.9.1 (Microsoft) for a range of alarm states. The fact that DERS was being used extensively by the end users for continuous infusions was vital to the study, as critical short half-life infusions, such as adrenaline/epinephrine, noradrenaline/norepinephrine, dobutamine, dopamine, and glycerol-trinitrate/nitroglycerine (GTN/NTG), could be clearly identified. In the study group, the DERS software was activated for infusions in 67.39% of cases. This percentage of infusions started from within the library can be expected to increase significantly in the near future with the ability to push datasets via wired and wireless systems to all pumps across a facility. This will allow for more complete infusion data, not just from ICUs but also from dispersed locations, including clinics and off-campus sites [7]. This will give more complete data on infusion types, generally, and on critical short half-life infusions and event logs, in particular.

The DERS in use on the study pumps allowed for the creation of distinct profiles; this involved free-text entry and it was required to group similar entities to create four distinct types of critical care units for the purposes of comparison and contrast: (1) PICU, (2) CICU, (3) GICU, and (4) NICU. As examples of similar entities, intensive care nurseries (ICNs) and special care baby units (SCBUs) were bundled into NICU, and CICUs were bundled with coronary care units (CCUs).

For the purposes of the study, the following critical short half-life infusions were selected: (1) adrenaline/epinephrine, (2) noradrenaline/norepinephrine, (3) dobutamine, and (4) dopamine. These critical short half-life infusions were selected due to their importance and extensive use in all areas of critical care, their potential for patient harm if any interruption to infusion is extended, and the fact that their plasma half-lives have been identified. Free-text entry from drug names, as well as generic terminology, and the differing languages present in the European region also required study grouping of the critical short half-life infusions into generic medication names (see Table 1). GTN/NTG was initially included in the study group, but once data interrogation began it was found that, while it had fairly extensive use in CICUs, its use did not extend across all critical care areas to a sufficient degree. It was therefore excluded from the study.

**Table 1.** Critical short half-life medications with plasma half-lives and sequelae to extended infusion interruption.

<table>
<thead>
<tr>
<th>Critical short half-life infusion</th>
<th>Plasma half-life in seconds</th>
<th>Possible impact of infusion interruption</th>
<th>Possible impact of postocclusion bolus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrenaline/epinephrine</td>
<td>180</td>
<td>Hemodynamic instability and severe hypotension</td>
<td>Hemodynamic instability and severe hypertension</td>
</tr>
<tr>
<td>Dobutamine</td>
<td>120</td>
<td>Hemodynamic instability and hypotension</td>
<td>Coronary artery contraction</td>
</tr>
<tr>
<td>Dopamine</td>
<td>60-120</td>
<td>Hemodynamic instability and hypotension</td>
<td>Coronary artery contraction</td>
</tr>
<tr>
<td>Noradrenaline/norepinephrine</td>
<td>120-180</td>
<td>Hemodynamic instability and severe hypotension</td>
<td>Coronary artery contraction</td>
</tr>
</tbody>
</table>
Infusion Device Types Included in the Study

The LVPs included in the study were CareFusion, BD Alaris Plus and Guardrails Plus general purpose and variable pressure pumps. The SPs included in the study were CareFusion, BD Alaris Plus and Guardrails Plus critical care and general hospital pumps. All of these pumps carry a common DERS, which applies soft and hard limits to all infusions delivered using the drug library. Soft Limits give the clinician a warning whenever a programmed administration rate or total dose is above or below an accepted dose range. This limit can be overridden by the clinician, thus allowing for clinical judgement to be employed, but only after an advisory message has been given by the DERS and acknowledged by the clinician. Hard limits require the clinician to reprogram a safe dose or rate. All soft and hard limit breaches are recorded by the pump. These flagged-up, near-miss events are recorded in a CQI reporter log, which can be downloaded directly from the pump or sent directly to the BD Alaris Communication Engine server for connected pumps. The pump’s event log is bundled with this CQI data.

The critical care and general hospital SPs include features designed to maintain infusion continuity and to promote steady plasma levels of infused medications. Both have Fast Start [8], which guarantees uptake of mechanical slack (ie, 95% take-up has been achieved by the best-performing pumps) [8], and a highly accurate volumetric purge function [9]. Both also carry a Back Off feature, which briefly draws back on the plunger if the occlusion alarm is activated, thereby reducing the size of any postocclusion bolus.

The critical care SP and variable pressure LVP also have in-line pressure monitoring, which is capable of detecting 1 mmHg increments of change in downstream pressure and is monitored via a diaphragm incorporated into the infusion line. This level of accuracy is important to this study as a sensitive, accurate, and rapid time to alarm, since any downstream occlusion is to be expected when these pumps are used for critical short half-life infusion delivery.

The international and independent medical device evaluator KLAS [10] has rated currently available commercial syringe drivers that have a particular focus on factors that reduce the risk of extravasation injury and harm to patients, and that help maintain infusion delivery continuity. These factors were as follows:

1. Shortening the time to alarm for occlusions.
2. Accurately monitoring in-line pressure as a proactive antietravasation measure.
3. Reducing delays of vasoactive medications at infusion start-up.
4. Decreasing inadvertent bolus following occlusion release.

The above technology would be expected to significantly improve steady-state infusion recovery times for critical short half-life infusions from any interruption to flow, where:

\[
\text{Total infusion recovery time} = \text{Time to alarm} + \text{Reaction time} + \text{Resolution (1)}
\]

This is the case, as time to alarm would be shortened by the presence of in-line pressure monitoring, and the resolution phase would be shortened by the Back Off and Fast Start features.

Taken from the above, the following is the case:

\[
\text{Total noninfusion time} = \text{Time to alarm} + \text{Reaction time (2)}
\]

This of course is also dependent on accurate and appropriate alarm setting [11]. The critical care SP and variable pressure LVP studied here also feature Auto-offset, which automatically sets the alarm pressure to a preset level above recorded line pressure and with Auto-set pressure, which will automatically set this limit at a predetermined time after the infusion begins. Both of these features are useful in detecting early changes related to vein infiltration and occlusion, and it is generally suggested that the Auto-offset be set at 30 mmHg above line pressure [11]. The reason for this is that high-set alarm pressures will delay the alarm even though the patient’s vein is already starting to be in distress or there is a hard occlusion, such as a clamp being inadvertently left closed (see Figure 1).

Reaction time is dependent on the clinician being alerted, prioritizing the alarm, responding, and resolving the issue. In many ways, this is the weak link in the chain for critical short half-life infusion maintenance, as even with pumps that carry technology capable of generating an appropriate time to alarm and with appropriately set alarms, resolution of the issue and re-establishment of the infusion is still dependent on the clinician responding to the alarm. Hence our desire to study reaction time in depth and to isolate it as a distinct event and issue for which solutions can be offered for extended and clinically unacceptable reaction times.

The pumps in the study group carry the same pattern of alarm priority across both LVPs and SPs. These generate both audible and visual alarms; therefore, proximity to the alarm is a key factor in the initial recognition of the alarm state. The process of prioritization is assisted by a high-medium-low-normal scheme (see Table 2).

Note that at this bedside level, there is no alarm differentiation between drug types (ie, a low-acuity medication alarm for occlusion would audibly and visually go off in the same way as a critical short half-life infusion would with the same issue). This is one aspect of the issue of alarm fatigue or “crying wolf” [12].
Figure 1. Time to alarm for dynamic alarm pressure setting and a fixed alarm setting for cannula-related and mechanical downstream occlusions.

Table 2. Alarm notifications by priority.

<table>
<thead>
<tr>
<th>Status</th>
<th>Pump beacon color</th>
<th>Alarm tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>High priority</td>
<td>Red</td>
<td>Rapid high pitch</td>
</tr>
<tr>
<td>Medium priority</td>
<td>Amber</td>
<td>Rapid low pitch</td>
</tr>
<tr>
<td>Low priority</td>
<td>Amber</td>
<td>Slow low pitch</td>
</tr>
<tr>
<td>Condition normal</td>
<td>Green</td>
<td>Nil</td>
</tr>
</tbody>
</table>

Database, Infusion Data Type, and Interface

The drug libraries of the study pumps have the capacity to hold 30 care areas or profiles and have the capacity for 3000 customizable drug types to be created in their respective libraries. As discussed above, the DERS has hard and soft limits for drug dosing and for patient weight, drug concentration, and drug boluses, as well as the ability to set specific occlusion alarm limits for each drug with the critical care and variable pressure pumps.

Any breach of these hard or soft limits is recorded, as well as total drug use across all pumps and compliance with drug library use in the CQI database. Data may be either manually loaded into these databases by a serial connection of pumps to client PCs or aggregated via a real-time infusion data collection service (ie, BD Alaris Communication Engine server software).

Study Procedure

The data were already patient anonymized, as no personal data such as hospital number, gender, name, data of birth, or other identifiable data were recorded or held on the pumps’ CQI logs or event logs. A further process of anonymization took place with no queries made for institution or facility. Only generic names of profiles or care areas remained.

Inclusion Criteria

Profiles or care units needed to be grouped into one of the following groups, in order to be included in the study: CICU, NICU, PICU, or GICU. For the specific investigation of critical short half-life infusion-interruption alarms, inclusion into the study required the presence of a clearly identifiable drug selection in one of the following groups: adrenaline/epinephrine, dobutamine, dopamine, or noradrenaline/norepinephrine.

Exclusion Criteria

Pump profiles outside of critical care areas were excluded, as the absence of critical short half-life infusions in these areas risked distorting the volume of alarms against which clinicians must identify and prioritize critical short half-life infusion-interruption alarms. A simple count of whole hospital infusion alarm types was, however, included for comparative value.

Data Manipulation and Interrogation

Data storage occurred via the CQI database—total data of approximately 60 GB—and the data were queried via SQL Server Management Studio v17.9.1 (Microsoft). Unique code was created for identification of specific pumps by count (see Multimedia Appendix 1, Table A1) and by significant alarm
conditions for all pumps, as well as by profile (see Multimedia Appendix 1, Table A2). In order to review the total alarm count by profile across a large cross section of pumps from differing facilities, it may be required to carry out the interrogation given in Multimedia Appendix 1, Table A3, to identify the profile types.

The nomenclature used as a profile name or label may differ across organizations for the same type of critical care unit; this was the case in our sample group, as it was spread across the European region. For example, Table 3 shows those that were grouped under NICU.

Table 3. Profile names grouped under the neonatal intensive care unit (NICU).

<table>
<thead>
<tr>
<th>Profile name</th>
<th>Profile ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal</td>
<td>11</td>
</tr>
<tr>
<td>Neonates &lt;2 kg</td>
<td>12</td>
</tr>
<tr>
<td>Neonates &lt;2 kg</td>
<td>13</td>
</tr>
<tr>
<td>NEOATLOGIE</td>
<td>3</td>
</tr>
<tr>
<td>NEONATLOGIE</td>
<td>4</td>
</tr>
<tr>
<td>NICU</td>
<td>14</td>
</tr>
<tr>
<td>NICU 3</td>
<td>15</td>
</tr>
<tr>
<td>NICU rule of 6</td>
<td>16</td>
</tr>
<tr>
<td>SCBU(^a)</td>
<td>79</td>
</tr>
<tr>
<td>SCN(^b)</td>
<td>21</td>
</tr>
</tbody>
</table>

\(^a\)SCBU: special care baby unit.
\(^b\)SCN: special care nursery.

This grouping gave the profile ID numbers to be used for queries. The code used to interrogate the number of downstream occlusions in NICU is given in Multimedia Appendix 1, Table A4, as an example.

Where the nomenclature of the specific profiles used in the dataset and drug library are known, as would be the case in a single facility querying its data, the code given in Multimedia Appendix 1, Table A5, may be utilized. For pump count and pump type for specific profiles, the code given in Multimedia Appendix 1, Table A6, may be utilized.

The gross number of undifferentiated infusions per day was calculated from the CQI reporter software, which gives both total infusions started over the time period investigated and a breakdown of those started from within the drug library using the DERS. For the investigation of the number of alarms occurring specifically for critical short half-life infusion alarms, the various codes used are given in Multimedia Appendix 1, Table A7.

A manual review of the data obtained from the above code to exclude repeat alarms completed the analysis to allow for only the inclusion of first-time reaction time to infusion-interruption states and alarms to be recorded in the results of the study for the following:

\[\text{Reaction time [in seconds]} = (\text{Time [hh:mm:ss] issue resolved and pump restarted}) - (\text{Time [hh:mm:ss] pump stopped and alarm triggered}) \]

### Results

Table 4 expresses the total and individual alarms and all infusion starts for each area of critical care and the whole hospital. From this, it can be identified that not only are the bulk of alarms seen in critical care, certainly due to the sheer number of infusions given in these areas, but there are also far more alarms per infusion in these areas too. Clearly the risk of alarm fatigue is greatest in these areas.

Infusion interruption of critical short half-life infusions was found to be a significant problem in all areas of the general critical care pump population with EOI and end of syringe events being noted, as well as a significant number of downstream (ie, vein and access) occlusion events. See Table 5 for a breakdown of critical short half-life infusion in the NICU by alarm type for each drug type. See Table 6 for a summary of restart times following alarms for the interruption of critical short half-life infusions.
Table 4. Infusion and alarm frequencies by care area and across the whole hospital between the study dates January 1, 2000, to September 29, 2017 (6482 days).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Frequency in critical care areas, n</th>
<th>Frequency in whole hospital, n</th>
<th>Alarm frequency, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>NICU&lt;sup&gt;a&lt;/sup&gt;</td>
<td>GICU&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Total infusion starts</td>
<td>187,441</td>
<td>30,527</td>
<td>45,756</td>
</tr>
<tr>
<td>Alarms per infusion, mean</td>
<td>4.50</td>
<td>3.71</td>
<td>4.36</td>
</tr>
<tr>
<td><strong>Alarm types</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total alarms</td>
<td>467,437</td>
<td>113,277</td>
<td>199,482</td>
</tr>
<tr>
<td>Flow error (ie, drip counter)</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Air-in-line accumulation exceeded</td>
<td>6975</td>
<td>N/A</td>
<td>6700</td>
</tr>
<tr>
<td>Air-in-line single bubble exceeded</td>
<td>17,100</td>
<td>N/A</td>
<td>12,724</td>
</tr>
<tr>
<td>Callback</td>
<td>134,754</td>
<td>22,396</td>
<td>68,037</td>
</tr>
<tr>
<td>Door open while infusing</td>
<td>14</td>
<td>N/A</td>
<td>14</td>
</tr>
<tr>
<td>Drive engage failure</td>
<td>10,850</td>
<td>5629</td>
<td>175</td>
</tr>
<tr>
<td>End of infusion</td>
<td>20,762</td>
<td>4650</td>
<td>849</td>
</tr>
<tr>
<td>Near end of infusion</td>
<td>107,236</td>
<td>10,695</td>
<td>49,517</td>
</tr>
<tr>
<td>Occlusion (downstream)</td>
<td>136,148</td>
<td>62,944</td>
<td>49,444</td>
</tr>
<tr>
<td>Occlusion (upstream)</td>
<td>14,602</td>
<td>N/A</td>
<td>11,710</td>
</tr>
<tr>
<td>Syringe disengaged</td>
<td>15,400</td>
<td>6738</td>
<td>290</td>
</tr>
<tr>
<td>End of syringe</td>
<td>3396</td>
<td>225</td>
<td>22</td>
</tr>
<tr>
<td>Infusions started from within the drug</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>library (ie, detectable as critical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>infusions)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>NICU: neonatal intensive care unit.
<sup>b</sup>GICU: general intensive care unit.
<sup>c</sup>CCU: coronary care unit.
<sup>d</sup>PICU: pediatric intensive care unit.
<sup>e</sup>N/A: not applicable.
Table 5. Critical short half-life infusions and alarms in the neonatal intensive care unit (NICU).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Frequency by drug type, n</th>
<th>Critical short half-life infusions and alarms for four drugs, n</th>
<th>Total infusion starts and alarms, n</th>
<th>Critical short half-life infusions and alarms for four drugs out of total infusion starts and alarms, %</th>
<th>Alarms for four drugs out of all critical short half-life infusion starts (N=3623), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total infusion starts</td>
<td>2509</td>
<td>96</td>
<td>58</td>
<td>960</td>
<td>3623</td>
</tr>
<tr>
<td>Alarm types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Callback</td>
<td>2776</td>
<td>15</td>
<td>28</td>
<td>459</td>
<td>3278</td>
</tr>
<tr>
<td>Drive engage failure</td>
<td>25</td>
<td>19</td>
<td>0</td>
<td>13</td>
<td>57</td>
</tr>
<tr>
<td>End of infusion</td>
<td>38</td>
<td>13</td>
<td>1</td>
<td>4</td>
<td>56</td>
</tr>
<tr>
<td>Near end of infusion</td>
<td>342</td>
<td>1</td>
<td>19</td>
<td>92</td>
<td>454</td>
</tr>
<tr>
<td>Occlusion (downstream)</td>
<td>267</td>
<td>114</td>
<td>12</td>
<td>120</td>
<td>513</td>
</tr>
<tr>
<td>Syringe disengaged</td>
<td>29</td>
<td>21</td>
<td>0</td>
<td>28</td>
<td>78</td>
</tr>
<tr>
<td>End of syringe</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

aDOB: dobutamine.
bDA: dopamine.
cAD: adrenaline.
dNAD: noradrenaline.
eN/A: not applicable.

Table 6. Time to restart following alarms for critical short half-life infusion interruptions.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Reaction times in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neonatal intensive care unit (NICU)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>58.73 (43.63)</td>
</tr>
<tr>
<td>Maximum</td>
<td>444.25</td>
</tr>
<tr>
<td>Minimum</td>
<td>4.00</td>
</tr>
<tr>
<td>Median</td>
<td>17.50</td>
</tr>
</tbody>
</table>

Discussion

Table 4 shows an initial analysis and a mean average of 4.50 alarms per infusion in the general critical care pump population. In the PICU, the alarms per infusion value was significantly above the mean average for all critical care areas, with 8.61 alarms per infusion. Both of these values are far above that seen in the general whole hospital pump population (ie, 1.39 alarms per infusion), which includes all pumps in critical care and all other medical-surgical units, clinics, oncology centers, and all other care areas.

In terms of possible strategies for the mitigation of alarm fatigue, it is of value to review the two parameters NEOI and EOI. NEOI is an indicator to the clinician that an infusion is close to completion or close to empty in terms of the syringe or bag. NEOI was responsible for almost 13% of all infusion alarms (see Table 4) and occurred in 57% of all infusion starts. This is a large volume of alarms, which may or may not be significant; as discussed earlier, the alarm for NEOI, as for all alarms, is not differentiated between critical short half-life infusions and noncritical medications. Indeed, in the case of an intermittent drug such as an antibiotic, the clinician may in fact need to run the infusion until they empty the bag or syringe in order to give a full dose; in this case, the NEOI alarm is truly a nuisance alarm.

A centralized system for infusion monitoring can overcome much of this issue of alarm differentiation, provided that the medications are clearly identified as critical short half-life infusions or noncritical and are easily located or mapped from display to geographical location. While the alarm may continue, the clinician would be able to decide on the appropriate level of response.

For EOI and end of syringe alarms, the question of rapid response versus standard response very much depends on the question of whether the medication is a critical short half-life infusion or not. EOI and end of syringe alarms account for only 1% and 0.17% of all infusion alarms, respectively (see Table 4), and would therefore not contribute overly to alarm fatigue. However, these alarms are very clearly analogous to the problem of “crying wolf” identified by Waterson et al [12], in that they
occur in a total of 1.7% of NICU critical short half-life infusions (see Table 5). For instance, they are rare events but potentially extremely hazardous and hard to detect and differentiate. This is the case, given the fact that EOI and end of syringe alarms are the same for all medications and are presented to the clinician among a large volume of alarms, even among the subgroup of critical short half-life infusions where 90.5% of these infusions had the noncritical alert Callback associated with them (see Table 5). Again, a centralized routed alarm management system with differentiated alarm tones and visual indicators for critical short half-life infusions would almost certainly be of value to reduce the “crying wolf” risk.

The ratio of downstream occlusion to infusion starts in the general pump population was 0.52; in the critical care areas it was 1.38, while in PICU and NICU it was 2.71 and 2.06, respectively (see Table 4). Given that extravasation injury remains the most common iatrogenic neonatal injury [13], it is clear that these occlusion alarms are not merely nuisance alarms and that prompt attendance to them, followed by inspection of the intravenous cannula site and appropriate action, may prevent worsening of any infiltration or extravasation injury [14]. They are urgent alarms and they are also as likely as critical short half-life interruption alarms to be missed among the white noise of the mass of alarms to which clinicians are exposed. Any centralized or distributed alarm system employed to ensure rapid and appropriate response must be capable of indicating clearly the particular problem to be addressed by the clinician. Such a system should also be capable of indicating parameters such as vein pressure to allow clinicians to become proactive in infusion management. This study indicates how capturing infusion data from right across the health care facility makes it possible to identify problematic areas of the hospital for potential alarm fatigue as well as particular issues of infusion and infusion line management in specific areas.

The mean and median response times to critical short half-life infusion interruption were generally within the half-lives of the selected medications (see Table 6). However, there was a high prevalence of outliers in terms of reaction times for all the critical short half-life infusions studied, with some response times being longer than 13 minutes or 6 times greater than the accepted plasma half-life of the concerned medication (see Figure 2).

Escalation of alarms from pumps that may be in isolation rooms or simply in geographically dispersed locations has been shown to be of value in one study where central infusion monitoring was introduced to an NICU along with environmental changes. A 56.25% reduction in measured infusion alarms was recorded in the study as well as a 31% reduction in reaction time to infusion alarms for critical short half-life infusions [15].

With the trend toward increasing numbers of satellite critical care and high-dependency beds, more isolation rooms, and increased numbers of chronic critical care patients, the number of complex intravenous infusion therapies given in dispersed locations is likely to increase. The need for strategies and technologies to manage the risk of alarm fatigue and prolonged interruption of critical short half-life infusions has been identified in this study, along with particular issues of infusion management that warrant further attention and appropriate interventions. Indeed, pump data have been previously used to help inform and support intervention planning, such as added training and resetting default pump alarm settings, by targeting specific problems such as downstream occlusions and specific care areas based on the data pulled from pumps [2]. These interventions only produced a 2% reduction in alarms per patient per day. It was felt that this limited impact was due to the fact that the project was limited in its flexibility, as “weekly adjustments to the process” were not realistic given the time required for data extraction and analysis; as well, the project was limited geographically, as the project could only be completed in one hospital unit [2].

This study lays out baseline data for possible benchmarking from units right across the hospital. It also has the flexibility and speed of application through established SQL (structured query language) query codes to allow for rapid and comprehensive data processing, allowing for rapid analysis of change pre- and post- any intervention. The accumulation of data can, of course, also be expected to increase with networked pumps, particularly if wireless pumps are deployed to dispersed locations within the hospital. Such an accumulation of data would allow for a deeper look into noncritical care areas where alarm management might benefit from an extension of alarm notification beyond central monitoring to clients in the form of tablet computers or other handheld devices.

One clear conclusion from this study is that the medication administration rights [16] should be extended to include right maintenance and ensured delivery continuity of critical short half-life infusions. There has been much useful focus in medication safety on ensuring correct and safe initial setup of intravenous medication infusions. However, without a focus on ensuring continuous and uninterrupted delivery of infusions, particularly in the case of short half-life medications, much of the effectiveness of these setup strategies is essentially redundant.
Figure 2. Restart reaction times after critical short half-life infusion interruption alarms in the neonatal intensive care unit (NICU): means (horizontal lines within the boxes), quartiles (upper and lower edges of the boxes), and outliers (red Xs). Also shown are indicative figures for consequences of prolonged interruption of critical short half-life infusions and the half-lives of critical medications in seconds. IV: intravenous.

Acknowledgments

This research was internally funded by the Medical Affairs Department of Becton, Dickinson and Company (BD), a manufacturer of a broad portfolio of health care products, which includes large volume infusion pumps and syringe pumps. The company is committed to enhancing patient, clinician, and environmental safety through scholarly scientific research.
Conflicts of Interest

All authors certify that they have no financial affiliations with, or involvement in, any organizations or entities with a financial interest, beyond their employment at Becton, Dickinson and Company (BD).

Multimedia Appendix 1

Microsoft SQL Server Management Studio v17.9.1 query codes. SQL: structured query language.

References


Abbreviations

AD: adrenaline
BD: Becton, Dickinson and Company
CCU: coronary care unit
CICU: coronary (adult) intensive care unit

http://humanfactors.jmir.org/2019/3/e14123/ JMIR Hum Factors 2019 | vol. 6 | iss. 3 | e14123 | p.12 (page number not for citation purposes)
CQI: continuous quality improvement
DA: dopamine
DERS: dose error reduction software
DOB: dobutamine
EOI: end of infusion
GICU: general (adult) intensive care unit
GTN/NTG: glycerol-trinitrate/nitroglycerine
ICN: intensive care nursery
ICU: intensive care unit
LVP: large volume pump
NAD: noradrenaline
NEOI: near end of infusion
NICU: neonatal intensive care unit
PICU: pediatric intensive care unit
SCBU: special care baby unit
SCN: special care nursery
SP: syringe pump

Edited by G Eysenbach; submitted 24.03.19; peer-reviewed by P DeLaurentis, M Cvach; comments to author 17.04.19; revised version received 22.05.19; accepted 19.06.19; published 12.08.19.

Please cite as:
Waterson J, Bedner A
Types and Frequency of Infusion Pump Alarms and Infusion-Interruption to Infusion-Recovery Times for Critical Short Half-Life Infusions: Retrospective Data Analysis
JMIR Hum Factors 2019;6(3):e14123
URL: http://humanfactors.jmir.org/2019/3/e14123/
doi:10.2196/14123
PMID:31407667

©James Waterson, Arkadiusz Bedner. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 12.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
Review

Social Cognitive Theories and Electronic Health Design: Scoping Review

Patricia Grace-Farfaglia, MSc, MBA, RD, PhD
Department of Nutritional Sciences, University of Connecticut, Waterbury, CT, United States

Abstract

Background: There are several social cognitive theories (SCTs) and models that support platform design in electronic health (eHealth) promotion trials. The rationale for this scoping review was to determine how social design features (informational aid, expressive support, gaming, and tailored content) are used to promote self-efficacy, engagement, knowledge, and behavior change.

Objective: This study aimed to review a broad spectrum of digital health interventions in the literature seeking trials that use SCTs for the design of eHealth applications.

Methods: The author conducted a systematic scoping review of 161 Web-based health interventions from published randomized clinical trials using 1 or more tools to address the social cognitive determinants in their website design from January 2006 to April 2016. An iterative approach was used in the selection of studies and data extraction. The studies were analyzed for quality and coded for type of social design features employed.

Results: Expressive interaction tools were found in 48.6% (54/111) of studies categorized as a strong recommendation by the Joanna Briggs Institute criteria. Overall, less than half of the studies addressed participant social support and motivational needs (43.8%). The vast majority of studies (100%) relied on the use of the Web for delivery of informational aid and tailored content for the individual participant (75.9%).

Conclusions: This review fills a research gap by linking social theory to Web strategy to improve the impact and sustainability of eHealth interventions. A Digital Health Intervention Model was developed to provide a framework to enhance future Web-based health intervention design and execution.

Introduction

Background

The aim of this scoping review was to review published clinical trials for evidence of social cognitive theory (SCT)-driven design of electronic health (eHealth) interventions. A secondary objective was to review the construct of social support as it applies to Web-based social interactions and recommend the inclusion of tools that increase media social presence to explore its value in the design and measurement of health behavior outcomes [1]. Finally, this paper offers an integrative theoretical perspective and framework for health researchers and designers to engage users to remain digitally connected for better health outcomes.

Social Theories and Electronic Health Design

Early in the internet era, Bandura envisioned that SCT would function as part of a self-regulatory delivery system for computer-assisted health interventions [2]. Bandura’s work demonstrated that learning takes place within a social context. A 2015 systematic review and meta-analysis of SCT-based interventions for patients with cancer for diet and physical activity (PA) found only 18 articles that met the inclusion
criteria of reporting 1 or more SCT constructs in the design [3]. The scarcity of theory-based design studies leaves a wide gap in our understanding of the Web-based experience factors contributing to health behavior change and maintenance [4]. Authors of a recent systematic review developed a taxonomy of 36 social media features and described their use in 134 studies, reporting that the majority reported positive effects including engagement, satisfaction, usefulness, social support, and behavior change (70%) [5]. However, the remainder of the papers reported no behavioral change (28%) or negative outcomes (2%). These findings suggest that social influence tools in digital behavior change interventions may produce unintended effects.

SCT is an integrated model of emergent interactive agency where personal factors and environmental events function as interacting dependent variables and operate as reciprocal factors predicting behavior. This theoretical framework has greater saliency today, as more consumers digitally track their health behaviors and are connected to external interactive guidance and social support [6]. In SCT, the constructs of self-observation, judgmental process, and self-reaction comprise a system of self-regulation of motivation and behavioral action. However, too often in the studies of Web-supported interventions reviewed for this paper, social interactions took second place to techniques that offered information, monitoring, and self-management tools [7].

SCTs, such as the health belief model, theory of reasoned action, and theory of planned behavior recognize the role of social support as an important determinant of health behavior. But it is essential that health researchers define the type of social support sought and received by the user to guide their study design and predict how social presence will perform within the digital medium before study implementation. The dimensions of Web-based social support have been classified as instrumental, socioemotional, and informational [8]. The digital environment gives ready access to asynchronous, simultaneous, and bidirectional social support.

A survey of 240 health-related websites rated the quality of social media tools and use of evidence-based theory for Web design [9]. Quality was determined by the presence of behavioral components, interactivity, and user-generated content in the design. Nearly half of the sites offered feedback, which consisted primarily of simple guidelines rather than tailored advice. The primary applications for content sharing were status updates, discussion forums, sharing success stories, sharing photos blogs, and comments. Overall, reviewers gave low marks to the sites as they lacked tools that promote theory-based behavior change.

Social Presence and Personalization Impact Behavior Change

Social presence, the perception of nonmediation, conveys the sensation of intimacy and immediacy in digital communication [10]. Studies comparing real-world face-to-face with digital discussions in health interventions have found that they promote adherence and behavior change [11,12]. As observed in studies of interactions of teens on Web-based social network, teens engage more fully when they cocreate content and develop Web-based self-identities through emojis, video, and other audiovisual materials [13] contributing to the feeling of being copresent with peers within a virtual environment [14]. Adult focus groups of users of a Web-based social network for health behavior change suggested to researchers that the addition of personalization options such as pictures, recipes, and status updates for social interaction and comparison are desirable options [15]. The quality of social presence is unique to each media channel; therefore, choice of medium has a direct effect on the depth of information processing and user motivation to take greater effort to process a message. Iterative health information sharing by social media users has resulted in many benefits including enhanced self-efficacy and healthy lifestyle adoption in multiple studies [16].

When faced with uncertainty, humans have higher information needs and will seek out trusted sources of information. Lee and Kvasny (2013) proposed that information richness and social presence of the Web experience satisfy the needs of an individual to obtain instrumental and expressive support [8]. Figure 1 displays a theoretical framework of how social media use and Web-based social support address these needs. This model illustrates that uncertainty is a consequence of a person’s self-appraisal of efficacy. The reduction of an uncomfortable state of uncertainty motivates the health consumer to seek out experts or Web-based peers.

As communities of caring have proliferated on the Web, so has the ability to access meaningful solutions for health problems. Social connections can be established based on similarities and differences important to the Web user and can be established through member profiles or disease-focused forums. An overview of the social media, social presence, and information richness characteristics of different websites is shown in Table 1 [8,17-19].

https://humanfactors.jmir.org/2019/3/e11544/
Figure 1. Social media and social support.

![Figure 1](https://humanfactors.jmir.org/2019/3/e11544/)

Table 1. Social media characteristics.

<table>
<thead>
<tr>
<th>Social media form</th>
<th>Information richness</th>
<th>Social presence</th>
<th>Social Support Scale 1-6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blogger or WordPress editable pages or wiki such as</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wikipedia</td>
<td>Essays and short reports via blogs</td>
<td>Blogs; narrative or storytelling</td>
<td>1-2</td>
</tr>
<tr>
<td>Google Docs</td>
<td>Essays and short reports via blogs</td>
<td>Blogs; narrative or storytelling</td>
<td>1-2</td>
</tr>
<tr>
<td>Audio sharing through Clyp or SoundCloud</td>
<td>Music and voice</td>
<td>Collective experience and content community</td>
<td>4</td>
</tr>
<tr>
<td><strong>Videos and collaborative spaces distributed through services</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pinterest</td>
<td>Music, voice, video, collaboration, and reviews</td>
<td>Collective experience and content community</td>
<td>4</td>
</tr>
<tr>
<td>YouTube</td>
<td>Music, voice, video, collaboration, and reviews</td>
<td>Collective experience and content community</td>
<td>4</td>
</tr>
<tr>
<td>Vimeo</td>
<td>Music, voice, video, collaboration, and reviews</td>
<td>Collective experience and content community</td>
<td>4</td>
</tr>
<tr>
<td>Viddy</td>
<td>Music, voice, video, collaboration, and reviews</td>
<td>Collective experience and content community</td>
<td>4</td>
</tr>
<tr>
<td><strong>Combined, multipurpose platforms that offer multiple media options</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook</td>
<td>Essays, short reports, and links to articles</td>
<td>Social network; pictures; videos; messaging; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>Ning</td>
<td>Essays, short reports, and links to articles</td>
<td>Social network; pictures; videos; messaging; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>Essays, short reports, and links to articles</td>
<td>Social network; pictures; videos; messaging; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>Skype</td>
<td>Essays, short reports, and links to articles</td>
<td>Social network; pictures; videos; messaging; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td><strong>Short-form text messaging and photo sharing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter</td>
<td>Texting: music, movies, books, maps, voice, high quality photos, and videos</td>
<td>Social network; pictures; immediacy; reach; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>Instagram</td>
<td>Texting: music, movies, books, maps, voice, high quality photos, and videos</td>
<td>Social network; pictures; immediacy; reach; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>Snapchat</td>
<td>Texting: music, movies, books, maps, voice, high quality photos, and videos</td>
<td>Social network; pictures; immediacy; reach; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td>Gtalk</td>
<td>Texting: music, movies, books, maps, voice, high quality photos, and videos</td>
<td>Social network; pictures; immediacy; reach; leaders and followers</td>
<td>5</td>
</tr>
<tr>
<td><strong>Mobile health (mHealth) apps</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eHealth record</td>
<td>Practice of public health and medicine through mobile devices</td>
<td>Collective experience; content community; information; motivation; support; remote monitoring; diagnostics and decision support</td>
<td>6</td>
</tr>
<tr>
<td>Social health</td>
<td>Practice of public health and medicine through mobile devices</td>
<td>Collective experience; content community; information; motivation; support; remote monitoring; diagnostics and decision support</td>
<td>6</td>
</tr>
<tr>
<td>Web-based health communities</td>
<td>Practice of public health and medicine through mobile devices</td>
<td>Collective experience; content community; information; motivation; support; remote monitoring; diagnostics and decision support</td>
<td>6</td>
</tr>
<tr>
<td>Fitness apps</td>
<td>Practice of public health and medicine through mobile devices</td>
<td>Collective experience; content community; information; motivation; support; remote monitoring; diagnostics and decision support</td>
<td>6</td>
</tr>
<tr>
<td>Personally controlled health management systems (PCHMS)</td>
<td>Practice of public health and medicine through mobile devices</td>
<td>Collective experience; content community; information; motivation; support; remote monitoring; diagnostics and decision support</td>
<td>6</td>
</tr>
<tr>
<td><strong>Gamification (serious health gaming)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiuser Dungeons (MUD)</td>
<td>Virtual social worlds with identities and collaborative content; challenge; competition; avatars</td>
<td>Collective experience; content community; shared emotional states; social network; social influence</td>
<td>6</td>
</tr>
<tr>
<td>Re-Mission2</td>
<td>Virtual social worlds with identities and collaborative content; challenge; competition; avatars</td>
<td>Collective experience; content community; shared emotional states; social network; social influence</td>
<td>6</td>
</tr>
<tr>
<td>DietBet</td>
<td>Virtual social worlds with identities and collaborative content; challenge; competition; avatars</td>
<td>Collective experience; content community; shared emotional states; social network; social influence</td>
<td>6</td>
</tr>
</tbody>
</table>
Scoping Review of Social Cognitive Theories in Electronic Health Design

This scoping review will summarize various SCTs and digital methods used within a wide spectrum of digital health interventions. During the planning of this paper, it was clear that early eHealth trials too often neglect to report a theoretical basis for their research design. Thus, the initial keyword search criteria yielded only 4 usable trials. With the 2010 publication of Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth (CONSORT-EHEALTH) guidelines, authors began specifying the mode of delivery and type of trial in the title of the paper, along with study details that would facilitate search retrieval [20].

Methods

Overview

Keywords were chosen to locate papers with the characteristics and tools of Web-enabled health intervention within a randomized controlled trial (RCT) or cohort study based on social cognitive or learning theories that informed study design. The researcher used a coding scheme based on the work of Lee and Kvasny with the addition of serious gaming to summarize the results so that research gaps and opportunities could be identified [8]. Using the guidelines of the Joanna Briggs Institute (JBI) manual for the levels of evidence and grades of recommendations and the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR), papers meeting the criteria for inclusion were graded and summarized [5,21].

Search Strategy

The University of Connecticut Library databases (PsychInfo, Cumulative Index to Nursing and Allied Health Literature [CINAHL], EBSCO), ProQuest, and PubMed were thoroughly searched using the search terms internet, Web, health, intervention, social presence, expressive aid, instrumental aid, information richness, and design for the years 2006 to 2016.

Intervention Criteria

Any health promotion or guidance designed to influence health behavior was based on SCT [6] or explicitly described and referenced any SCT component (such as self-efficacy):

- Comparator: any parallel control group.
- Study design: RCTs or cohort.

A total of 399 studies were identified by keyword search alone. Using CrossRef citations for the CONSORT-EHEALTH guidelines, 146 additional studies were identified as of April 2016. A total of 161 studies met the RCTs risk of bias and study quality [22,23]. Multimedia Appendix 1 outlines the PRISMA flowchart of the study selection process.

The trials retrieved were coded by theoretical framework, presence of informational and expressive support in the study design, and use of tailored content. Identification and coding of the theoretical framework was hindered by the need to cross-reference previously published papers identified by the authors as the basis for their protocol or multiple papers from the same study data. If the study design was consistent with the outcomes of that framework, it was retained. The most common model for tailoring the Web intervention was social-cognitive learning theory (N=106). Several studies identified therapeutic frameworks adapted for the Web, such as internet cognitive behavioral therapy (iCBT) [24]. The use of the iChange model for health behavior which used wearable measurement tools is a growing trend to facilitate lifestyle change by giving feedback to the participant [25-27]. Gamification, the use of games used seriously (GUS), and fun theory were successfully used to foster participation in 4 studies, and a recent meta-analysis of trials using games for health behavior change reported significant positive outcomes in 9 out of the 10 studies meeting their criteria [28-31]. The use of fun theory was uniquely used in serious video gaming and physical therapy where adherence to the daily practice of mundane tasks is fostered by a challenge and excitement [32,33].

A qualitative analysis of this literature was undertaken because of the diversity in methodologies. The highest quality evidence followed the CONSORT-EHEALTH or TREND guidelines, although a few quality trials predated the guidelines, and recent meta-analyses were included in the discussion [20,34,35].

Results

Overview

The literature search retrieved 447 results from PubMed (n=296), University of Connecticut Library databases (PsychInfo, EBSCO, and CINAHL; n=72), and ProQuest (n=79). After removal of duplicates (n=48), 399 titles and abstracts were screened, 215 full-text articles were reviewed, 4 were excluded as additional qualitative papers from a mixed-methods design trial, and 161 articles met the selection criteria (Multimedia Appendix 1). Upon further inspection, there were 9 papers that were overlapping manuscripts from the same dataset which brought the final total to 152 studies. Studies were coded by the author for theoretical framework, tools (informational aid, expressive aid, and gaming), and content that was tailored to each user (Multimedia Appendix 2 [27,36-196]). From this process, the author evaluated different theoretical frameworks and the use of Web intervention tools and environments to promote desired health outcomes. The JBI levels of evidence and grades of recommendations were used to evaluate the studies [187].

Risk Preference and Goal Setting

Setting personal health goals and evaluating one’s ability to attain them is an essential part of chronic disease management. It has been observed in behavioral economics that loss aversion is a powerful motivator, and this trait is important in setting personal goals. It has been observed that obese individuals are more likely to be risk-seeking, rather than risk-averse, when making decisions that offer uncertain options, which is a tendency related to impulsivity. In Prospect Theory, this is known as risk preference [188,189]. One study observed participant risk taking behavior by allowing participants to bet on the outcome of a weight loss challenge on a commercial website, DietBet [190]. The researchers theorized that offering
frequent and small incentives on a social gaming site would influence players to lose weight. Members placed a bet and joined a social game where they wagered that they could achieve the goal of losing 4% of their total body weight and reported their progress to other participants. The financial and social influencers were effective in supporting weight loss as measured by self-reported weight, bets placed, frequency of social interactions, and weigh-in reports on Facebook. A month-long study of the effect of financial incentives within a social GUS environment reported that incentives coupled with social influence promoted greater weight loss [190]. In this way, social gaming facilitates the development of intrinsic motivation through the gratification of entertainment, challenge, and competition from game playing and harnessing the power of others [191,192]. There have been conflicting reports of the effectiveness of extrinsic incentives and/or penalties to promote health behaviors, depending on the target user region, gender, race, and income [193,194].

Goal setting is an important factor in achieving self-regulatory health behavior. Bandura and associates assigned obese subjects to goal conditions in which they either tracked eating behavior or set subgoals for reducing portions [195]. Within the goal-setting conditions, subjects adopted either weekly or proximal goals for each of 4 time periods during each day. The results demonstrated that setting a higher standard for goals and the adoption of proximal goals resulted in greater weight loss. As demonstrated in the Pagoto et al (2014) study, predetermined proximal goals in combination with Web-based social support improve adherence, support for others, and self-regulation. Research suggests that breaking down goals into subgoals may influence subsequent goal pursuit by reducing goal pursuit because competing subgoals may be perceived as complementary and become substitutes for one another [196].

A Web-based randomized controlled weight loss intervention combined PA and nutrition interventions over a 12-month period [197]. Strategies were guided by SCT theory and included goal setting, self-monitoring, and social support. Participants in the experimental group achieved positive health behavior change (mean z score=1.34 [P<.001] SD units). An interactive Web-based program was developed to set goals relative to the participant’s initial stage of change, revise goals frequently, track behaviors, and deliver graphical feedback. This study demonstrated that interventions can successfully target multiple behaviors simultaneously.

**User-Centered Design**

Design objectives for building a community of support for health and wellness should include Web-enabled interaction between the environment and individual choices, which comprise essential components of Bandura’s model of social determinism [198]. Bandura maintained that not only does the environment impact behavior, but human behavior influences the environment. Therefore, health sites should offer tools that support the affective, cohesive, interactive, and social presence needs of the site visitor to increase program engagement [199,200]. An example of a Web-enabled clinical trial, the Enabling Mothers to Prevent Pediatric Obesity Through Web-Based Education and Reciprocal Determinism (EMPOWER) study, focused on encouraging mothers to make changes in the home environment, develop coping skills, form positive expectations, and build self-efficacy [185,201]. The EMPOWER program was delivered in short audiovisual educational presentations, goal planning exercises, tracking worksheets, and a discussion board. Process evaluation data were collected after each session using telephone counseling and Web-based surveys. As predicted by SCT, the intervention resulted in significant increases in child PA, fruit and vegetable consumption, and sugar-free beverage consumption compared with an informational approach. Another study tested the effects of social presence cues (2 staring eyes) on the activation of health-related goals within an ecommerce site [202]. The analysis yielded significant main effects for social presence ($F_{1,218}=5.89$; $P=.016$; $d=0.32$) and health goal activation ($F_{1,218}=4.11; P=.04; d=0.27$) on the selection of healthier menu choices compared with the control condition. The combined effects of social presence cues and health-related goal activation produced greater effects on food choices when activated at the same time. Social presence was also associated with the participant perception of success in self-regulatory behavior.

User-centered website design has been successfully used in a chronic disease intervention for patients with type 2 diabetes [203]. In this study, focus group members requested features that included personalized information about their health status, quantified self-tracking tools to monitor progress, and online forums to share their personal experiences. **Personally controlled health management systems** (PCHMSs) have been adopted in many health care organizations to give patients better ways to manage their health. Future studies should identify the best PCHMS tools and evaluate their effectiveness for disease management and prevention. A study on the use of the social and self-reflective features of a PCHMS was designed to support physical and emotional well-being, and frequent use of PCHMS was associated with help-seeking behaviors and increased health care utilization [204]. Although PCHMS tools are common in health care, only 1 trial was located in this search.

In the Self-Help, Exercise and Diet using Information Technology (SHED-IT) study, tailored health communications were specifically designed to address the desired health behavior of adult males and various predictor variables [205,206]. Adherence to the goal setting ($\beta=-0.3\ 95\%\ CI\ -0.6\ to\ -0.1;\ P=.01$) and volume of SCT tracking tasks completed ($\beta=-0.2\ 95\%\ CI\ -0.4\ to\ -0.0;\ P=.03$) independently predicted weight loss [207]. Message strategies best matched to individual health-related goals increased the impact of functional support. In a meta-analysis of 88 papers on computer health interventions, the use of dynamically tailored interventions gained efficacy over time [208].

**Online Health Communities and Social Networking**

A small feasibility trial used a social networking site (Facebook) to promote PA among low active teens who reported medium-to-large changes in PA as measured by accelerometer and self-report [209]. An RCT of college students evaluated the efficacy of a Web-based PA intervention that combined information, self-monitoring, and Web-based social networking strategies in comparison with an instruction-only control [210].
Participants were invited to the Internet Support for Healthy Associations Promoting Exercise (INSHAPE) study website to complete Web-based surveys on their perceived social support for PA (informational, esteem, and companionship subscales), and the Facebook Intensity Scale, a measure of engagement with social networking. The researcher observed and recorded Facebook interactions during the intervention, including comments, discussion forum posts, and affective responses to the comments of others (like button). The participation rates in this study were higher than in other published studies. The main effects from the analysis were PA time, esteem, and companionship social support. The authors concluded that real and virtual social connections should be used for group assignment and to match people by profile to encourage PA in future studies.

A mixed-methods study of SparkPeople®, a large online weight loss community, was undertaken to determine how social media frequency predicted perception of Web-based social support for weight loss [211]. The first phase of the study surveyed members for their experience with social support within the community using qualitative analysis of responses to open-ended questions (n=193). Survey respondents were frequent users of email, blogs, and forum discussions. The uses and gratifications for Web-based social support that emerged were informational, emotional, instrumental, appraisal, and network support [212]. The quantitative analysis examined the factor structure for social support from the 7 social media use items (n=187). Principal components analysis of social support items proved to be a 1-dimensional social media variable. Social media was a significant predictor of encouragement support (odds ratio [OR] 4.8, 95% CI 1.8 to 12.8; P<.001), but not for information or shared experiences support. The authors concluded that members sought encouragement and motivational support for other members, and finding these gratifications was a key determinant to behavioral modification persistence.

The subthemes identified in the study by Hwang et al (2010) included accountability, friendly competition, and humor for a weight loss community [212]. These themes were important motivators in a social media campaign that challenged friends and family to do 1-min core strengthening exercises each day and broadcast individual progress on Twitter (#PlankADay) and other social media platforms [213]. Starting out as a friendly challenge between associates, the researchers soon recognized that the message had a spreading network effect. Seizing on this opportunity for research, an Institutional Review Board review was requested and approval was given to begin data collection. The observations of this study are consistent with a framing effect hypothesis of positive messages under a low efficacy condition [214]. The #PlankADay message attracted the attention of people who desired core strength but did not take the time to work at gaining core strength. Positive messages lead to more effortful processing of the message. As the challenge was framed in a positive way and tweets contained humor, it is weighed more heavily than a negative message.

**Web-Based Lifestyle Coaching Supports Self-Efficacy**

The 4 major sources of self-efficacy beliefs can be transferred to a Web-enabled system: enactive mastery experiences, modeling, social persuasion, and psychological states [215]. A qualitative study of veterans in the Veterans Administration (VA) health system who were cancer survivors revealed many barriers as well as facilitators to their weight management goals [216]. The focus group identified wellness facilitators to boost self-efficacy such as information about their disease, being held accountable for their behaviors, motivational support from others, workout partners, and the ability to visualize healthy changes. Web-based tools enabled the veterans to believe that they can achieve their health goals. The VA health care system pioneered the use of telemedicine and in 1 study compared the efficacy of home TeleHealth monitoring for diabetes management with usual care plus monthly phone calls [217]. Although both groups showed improvement in blood glucose levels, the TeleHealth program results were superior. The TeleHealth program had the advantage of accountability and feedback for behaviors, whereas the monthly phone call with the nurse practitioner provides social presence and feedback. The addition of Web-enabled tools provides the social support that is often lacking in the lives of aging veterans.

Motivational interviewing (MI) is a client-centered counseling approach for eliciting behavior change. A collaborative partnership is developed between coach and client, allowing the participant to discuss goals from the previous week, and problem-solving tools to make goal revisions needed. In a review of Web-based interventions for type 2 diabetes, positive outcomes were associated with using barrier identification, problem solving, and self-monitoring techniques [218]. A randomized, controlled Web-based intervention for depression and stress offered problem solving therapy via email [219]. The goal was to reduce stress and anxiety by providing feedback on exercises that taught problem solving in a structured way. Recently, the use of avatars to deliver health coaching over a Web app was tested, but findings suggest that participants do not readily develop social relationships with avatars and this tool contributed little to the effects of the intervention [220,221].

Enactive learning occurs with a participant observing their own progress through self-monitoring tools, which has been shown to fortify the participant’s self-efficacy beliefs [222]. Bandura suggested that the self-regulation of behavior must be measured against the difficulty of individual obstacles [6]. Over time, the use of Web-enabled tools to monitor diet intake, PA, and other health metrics will provide a rich database to study the effect of past experience, goal setting, and health outcomes. Data from a Web-enabled weight loss intervention for African Americans suggest that early and frequent use of Web-based self-monitoring tools predicts greater weight loss [223]. Web monitoring facilitates positive self-talk by providing evidence of success.

**Social Network Adoption and Participation**

The recruitment and retention of network members to support health behavior adoption is another area of research. Social networking is a Web-enabled process that draws people and organizations together. Peers with experiential similarity offer support [224]. In an RCT, researchers sought to study the reciprocal learning process (social proof, verbal persuasion, self-monitoring, and frequent feedback) between peers in a
Web-based social media intervention [225]. The researchers issued a daily wellness challenge by text or email to 1503 participants, using a small steps approach to both the experimental and control groups. Participants in the intervention group who had access to the well-being challenges were encouraged to recruit others from their social network to join. The measures included the Individual-level Well-Being Assessment and Scoring Method (IWBS), the Interpersonal Support Evaluation List, and intensity of site use [226,227]. After 30 days, the IWBS was significantly greater in the social intervention group than nonsocial participants (µ=9.4 intervention vs 7.0 control group; P<.02). Intensity of use was a positive predictor of well-being.

Implementation of a Theory-Based Framework for Digital Health Intervention

The Digital Health Intervention model (DHIM), shown in Figure 2, integrates theoretical frameworks discussed in this paper and proposes a model of Web-based health promotion based on social cognitive, health behavior, and prospect theories. The DHIM could be applied in various health promotion settings from a worksite wellness to patient-centered medical homes. Researchers could use the results of a health screening questionnaire to populate a risk profile to generate a visual dashboard or navigator. Each contact with the site should have a social presence that recognizes the user as a unique individual and interactions should give the impression that the system is built to help them specifically. As people are slow to adopt new technology when ease of use and perceived utility do not meet user expectations, the technology should be tested for user-friendliness [228]. A confusing interface reduces effectiveness for disease management interventions [229].

Coaching can be delivered virtually through delivery of tailored messages, or through phone and chat synchronous communication with call center personnel. An interactive page with games and peer-to-peer connectivity gives the participant support for making behavioral change and opportunities to practice. Plans and goals become part of the dashboard profile available for the member and coach to review. For integrated personal health record systems and patient portals, email access for provider communication is important [230]. The user can seek instrumental aid for informational needs in the form of a video and library database. The development of an intentionally designed social network to support the health goals of the user supplies both instrumental and expressive aid through facilitating self-disclosure and emotional support exchanges between peers [231]. But for stigmatized health issues, such as HIV, informational needs and expressive support may best be met in anonymity rather than within a social networking environment [232]. The Web can enable those with mental health to come out of the shadows and seek treatment as demonstrated by the Web-based program, Considering Professional Help, designed to encourage veterans to seek mental health care [233].

Many health-related networks are imbedded with tools that allow the user to upload data from glucose meters, diet trackers, and exercise monitors. These health networks have been shown to positively impact the adoption of health behaviors [19,186]. Behavioral self-monitoring by recording, reporting, and revising action plans are significant predictors of goal attainment, but only when the participant believes that the health goal has motivational value [234]. Therefore, the reward value of the goal is an important factor in motivating behavior change [235].

Figure 2. Digital Health Intervention Model. A comprehensive health intervention program integrating social cognitive principles in a web-enabled pathway.
Discussion

Principal Findings

All of the studies reviewed used multimedia informational aids in their health management strategy. Expressive aid was found in 48.6% (54/111) of studies categorized as a strong recommendation. The use of targeted expressive aids was seen in 37.5% (15/40) of the lower quality trials. Only 4 studies included all study categories of tools (informational, expressive, gaming, and tailored content). Serious gaming interventions study design were all quality rated as Level 1 A. Higher quality interventions were more likely to employ 3 or more categories of behavior change techniques within their study designs. Owing to the diversity of health behaviors and treatment protocols analyzed in this review, it is not possible to follow-up with a systematic review of common design elements and their effectiveness at this time.

In the studies reviewed, the relative frequency of Web activities to deliver health intervention universally included informational aids within their design, but 1/4 of them did not tailor or personalize content. Almost half of the interventions did not offer tools to obtain expressive aid, and only 5 of the trials employed serious gaming. This pattern suggests that many health researchers believe that eHealth is primarily an evidence-based informational tool for acquisition and learning of specific health-related skills rather than a dose of social influence to encourage health behavior adoption [236].

Limitations

Many of the studies rejected in this search process were excluded because the authors failed to identify the a priori theoretical framework underlying the design strategy or referenced protocols from previous papers which were not theory driven. The risk of selection and publication bias is greater because of the lack of consensus as to publication guidelines that included theory in design considerations during this period. The Cochrane Database of Reviews recently withdrew the publication of a protocol to assess serious gaming health intervention studies with no explanation [237]. Publication guidelines are needed to improve the reporting of clinical trials using serious gaming for future systematic review and analysis. Few studies addressed how their Web-enabled protocol addressed the Health Insurance Portability and Accountability Act security rules or whether privacy was a concern for participants. The potential for innovative technologies such as virtual reality or artificial intelligence to individualize health care is great, but threats to personal health information security may produce distrust among users and health professionals. A systematic review of the literature on website design and interactivity concluded that security elements have an important impact on Web-based health information seeking [238]. These design elements are important patient-reported measures associated with research involvement and should be included in the design of future studies [239]. There is a need for theoretically driven, continuous cohort studies to assess the sustainability of user engagement with tailored digital tools for chronic disease management [240].

Recommendations

Researchers should define the theory or theories that guide their choices of interventions as well as the desired behaviors that lead to health outcomes. More research is needed to identify the conditions under which media richness and social presence enhance message processing. Creating an environment where social presence is part of the research design may lead to better study retention and greater understanding of Web-based peer-to-peer support. Researchers should consider the addition of a run-in-time phase to their intervention protocols where participant observations can establish the ecological validity of the environment before the intervention. Interactions with social agents, whether human or artificial, are the ultimate tools of social cognition.

Conclusions

SCTs provide a framework for design, implementation, and evaluation of health intervention programs and has been successfully used in several interventions presented in this paper. Creating Web-based environments where social presence and information richness are used as part of the overall strategy has several theoretical advantages. Web-based health interventions have the potential to act as sticky media that sustains the pursuit of desired diet and exercise behaviors long after the initial study phase is over. Using SCT to design digital interventions has been shown to have positive outcomes for weight control, PA, diabetes, mental health, nutrition, and wellness behaviors. The ultimate challenge for health practitioners is to integrate Web-enabled health communication into real-world health care.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The author is solely responsible for the research, review, and content of this manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.

[PDF File (Adobe PDF File), 89KB - humanfactors_v6i3e11544_app1.pdf]
Multimedia Appendix 2
Study characteristics. Internet-based intervention characteristics.

References


58. Carter MC, Burley VJ, Nykjaer C, Cade JE. Mental fitnes... full text] [doi: 10.2196/jmir.3402] [Medline: 25665498]


Abbreviations

CINAHL: Cumulative Index to Nursing and Allied Health Literature

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications

DHIM: Digital Health Intervention Model

eHealth: electronic health

EMPOWER: Enabling Mothers to Prevent Pediatric Obesity Through Web-Based Education and Reciprocal Determinism

https://humanfactors.jmir.org/2019/3/e11544/
GUS: games used seriously
iCBT: internet cognitive behavioral therapy
INSHAPE: Internet Support for Healthy Associations Promoting Exercise
IWBS: Individual-level Well-Being Assessment and Scoring Method
JBI: Joanna Briggs Institute
mHealth: mobile health
MI: motivational interviewing
PA: physical activity
PCHMS: Personally Controlled Health Management Systems
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
RCT: randomized controlled trial
SCT: social cognitive theory
SHED-IT: Self-Help, Exercise and Diet using Information Technology
VA: Veterans Administration

Edited by G Eysenbach; submitted 10.07.18; peer-reviewed by H Kondylakis, C Short, S Zheng, J Apolíndorio-Hagen; comments to author 30.08.18; revised version received 24.11.18; accepted 31.03.19; published 19.07.19.

Please cite as:
Grace-Farfaglia P
Social Cognitive Theories and Electronic Health Design: Scoping Review
JMIR Hum Factors 2019;6(3):e11544
URL: https://humanfactors.jmir.org/2019/3/e11544/
doi:10.2196/11544
PMID:31325290

©Patricia Grace-Farfaglia. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 19.07.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
Postinjury Complications: Retrospective Study of Causative Factors

Elizabeth Warnack¹, MD; Hersch Leon Pachter¹, MD; Beatrix Choi¹, MD; Charles DiMaggio¹, PhD; Spiros Frangos¹, MD, MPH; Michael Klein¹, MD; Marko Bukur¹, MD

NYU Langone Health/Bellevue Hospital Center, New York, NY, United States

Corresponding Author:
Elizabeth Warnack, MD
NYU Langone Health/Bellevue Hospital Center
462 First Avenue
Fifteenth Floor
New York, NY, 10016
United States
Phone: 1 (212) 263 2225
Email: elizabeth.warnack@nyumc.org

Abstract

Background: Injury care involves the complex interaction of patient, physician, and environment that impacts patient complications, level of harm, and failure to rescue (FTR). FTR represents the likelihood of a hospital to be unable to rescue patients from death after in-hospital complications.

Objective: This study aimed to hypothesize that error type and number of errors contribute to increased level of harm and FTR.

Methods: Patient information was abstracted from weekly trauma performance improvement (PI) records (from January 1, 2016, to July 19, 2017), where trauma surgeons determined the level of harm and identified the factors associated with complications. Level of harm was determined by definitions set forth by the Agency for Healthcare Research and Quality. Logistic regression was used to determine the impact of individual factors on FTR and level of harm, controlling for age, gender, Charlson score, injury severity score (ISS), error (in diagnosis, technique, or judgment), delay (in diagnosis or intervention), and need for surgery.

Results: A total of 2216 trauma patients presented during the study period. Of 2216 patients, 224 (224/2216, 10.10 %) had complications reported at PI meetings; of these, 31 patients (31/224, 13.8 %) had FTR. PI patients were more likely to be older (mean age 51.3 years, SE 1.58, vs 46.5 years, SE 0.51; \( P = .008 \)) and have higher ISS (median 22 vs 8; \( P < .001 \)), compared with patients without complications. Physician-attributable errors (odds ratio [OR] 2.82; \( P = .001 \)), most commonly errors in technique, and nature of injury (OR 1.91; \( P = .01 \)) were associated with higher levels of harm, whereas delays in diagnosis or intervention were not. Each additional factor involved increased level of harm (OR 2.09; \( P < .001 \)) and nearly doubled likelihood of FTR (OR 1.95; \( P = .01 \)).

Conclusions: Physician-attributable errors in diagnosis, technique, or judgment are more strongly correlated with harm than delays in diagnosis and intervention. Increasing number of errors identified in patient care correlates with an increasing level of harm and FTR.

(KEYWORDS
failure to rescue; error; harm; trauma

Introduction

Background
Performance improvement (PI) is a key component of trauma center operations and centers on increasing patient safety through reduction of harm and iatrogenic error. According to the National Coordination Council for Medication Error Reporting and Prevention, harm is defined as “impairment of the physical, emotional, or psychological function or structure of the body and/or pain resulting therefrom.” [1]. The Institute of Medicine estimates that as many as 98,000 patients die in hospitals as a result of preventable medical errors each year [2].
Before this report, there was little uniformity in classifying and reporting postsurgical complications and levels of harm [3]. In 2009, the World Health Organization defined health care–associated harm as harm arising from or associated with provision of health care [4]. This distinguished health care–associated harm, which was potentially preventable, from harm related to underlying patient disease. The World Health Organization proposed a 1- to 5-point harm scale, ranging from no harm to death and including mild, moderate, and severe levels of harm; this was meant to standardize these definitions for safety and quality reporting. In 2010, the Agency for Healthcare Research and Quality (AHRQ) added the duration of harm to this 1- to 5-point scale, with permanent harm defined as harm with lasting effect of 1 year or greater and temporary harm defined as having effects lasting less than 1 year [5]. By defining the levels of harm, hospitals can better classify patient and provider errors that contribute to poor outcomes. Historically, hospital quality metrics included adverse occurrence rate and mortality rate. Failure to rescue (FTR) is an evolving quality metric. FTR represents the likelihood of a hospital to be unable to rescue patients from death after in-hospital complications [6]. As a measure of hospital response to complications, FTR has been studied for patients undergoing major elective surgeries [6,7] and has recently been applied to the trauma setting [8]. Studies have shown that FTR is a better marker for hospital quality than mortality rate or complication rate alone [6,7], and it has been shown to be the primary driver of differences in hospital quality for trauma patients [8]. Trauma centers with low overall patient mortality are more successful at rescuing patients who experience complications [9]. Previous studies in the trauma literature have explored the association between the error type and likelihood of posttrauma mortality [7,9-12]. However, these investigations have not focused on the examination and classification of other posttraumatic complications or potential errors. There are few studies in the literature that categorize the level of harm from a given complication using the new AHRQ guidelines. In addition, few studies have explored the effect of type and the number of human errors on posttraumatic complications, likelihood of FTR, and levels of patient harm using a standardized scale. Our study of specific human errors and their effect on patient complications provides a unique contribution to the existing literature on FTR.

Methods

Patient Population

This was a retrospective study of trauma patients at Bellevue Hospital Center (BHC) in New York City. BHC is a large academic public hospital that is affiliated with NYU School of Medicine. The mean number of trauma patients admitted per year at BHC is 1500, with a 90:10 ratio of blunt to penetrating trauma [13]. Hospital catchment includes Manhattan and Western Brooklyn.

Data Collection

Patient demographic information, including age, gender, ethnicity, insurance status, and physiologic information such as systolic blood pressure, Glasgow Coma Score (GCS), and injury severity score (ISS) on admission, was abstracted from the BHC trauma registry from January 1, 2016, to July 19, 2017. In addition, data were abstracted from weekly PI records over the given period.

Identification of Harm Events

To assist with identification of as many complications/systems issues as possible, our program employs a PI coordinator who is an experienced physician extender, attends all trauma morning reports, and participates in daily walk rounds with the trauma service. We feel this allows us to capture adverse events in a timely fashion and ensure that they are reported to our PI meetings. All trauma surgeons in the department attend PI meetings and come to a group consensus for factors that contributed to a given complication. The discussions involve the entire trauma team but are led by the Trauma Medical Director. The Trauma Medical Director has taken a course focused on standardized PI Trauma Outcomes and Performance Improvement Course (TOPIC). The TOPIC incorporates the standardized definitions set forth by the AHRQ into its taxonomy classification.

The most commonly encountered complications in our institution are listed in our PI form, including deep vein thrombosis, pneumonia, chest tube–related complications, iatrogenic injury, death, and missed injury (Multimedia Appendix 1). A total of 7 factors contributing to these complications were examined, including delay in treatment, delay in intervention, error in diagnosis, error in technique, error in judgment, and nature of injury. Nature of injury was selected by attending surgeons if a patient’s underlying disease process (ie, severe medical comorbidities) or severity of injury contributed to a given complication. Multiple factors could be recorded for each complication. For patients who experienced multiple complications, only the most severe level of harm and its contributing factors were recorded. Physicians were also asked to rate the level of harm for each complication, using the standardized definitions set forth by the AHRQ (Multimedia Appendix 1) [5]. For simplicity of analysis, these ratings were then recoded into a 1- to 5-point scale, representing no harm, mild harm, moderate harm, severe harm, and unanticipated death.

Identification of Harm Events
Statistical Analysis

Factors associated with FTR and level of harm were modeled using logistic regression, controlling for age, gender, Charlson score [14], ISS [15], need for surgery (ie, if patient required a surgery for trauma on index admission), error (in diagnosis, technique, or judgment), delay (in diagnosis or intervention), nature of injury, and total number of factors. Ordinal logistic regression was used to assess factors contributing to the increasing level of harm. Lipsitz goodness-of-fit test was performed for the ordinal level of harm regression, and Hosmer-Lemeshow test and receiver operating curves were performed for the FTR regression. Chi-square analysis was used to compare demographic and physiologic characteristics of patients presented at PI meetings with all other trauma patients over the given period. For quantitative variables, the Wilcoxon rank sum test was used to compare values. \( P \) values less than .05 were considered significant. All statistical analyses were performed using SPSS Statistics version 23 (IBM Corporation). This study received institutional review board approval.

Results

Patient Population

A total of 2216 trauma admissions presented during the study period. Of 2216 patients, 224 (224/2216, 10.10%) were presented at PI meetings. Of these, 31 patients (31/224, 13.8%) identified as FTR. Of the patients with complications, 81 (81/224, 36.1%) died during their admission. Of these mortalities, 52 (52/81, 64%) patients were classified as anticipated mortalities without opportunity for improvement (OFI), 12 patients (12/81, 14%) were classified as unanticipated mortalities with OFI, and 17 patients were classified as anticipated mortalities with OFI (17/81, 20%; Table 1).

Table 1. Complications for performance improvement patients (N=224).

<table>
<thead>
<tr>
<th>Types of complication</th>
<th>Patients, n (%)</th>
<th>Number of deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abscess</td>
<td>3 (1.3)</td>
<td>0</td>
</tr>
<tr>
<td>Deep vein thrombosis</td>
<td>17 (7.6)</td>
<td>9</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>5 (2.2)</td>
<td>2</td>
</tr>
<tr>
<td>Clostridium difficile</td>
<td>3 (1.3)</td>
<td>1</td>
</tr>
<tr>
<td>Postoperative bleeding</td>
<td>2 (0.9)</td>
<td>0</td>
</tr>
<tr>
<td>Unplanned surgery</td>
<td>5 (2.2)</td>
<td>1</td>
</tr>
<tr>
<td>Chest tube</td>
<td>12 (5.4)</td>
<td>0</td>
</tr>
<tr>
<td>Iatrogenic injury</td>
<td>6 (2.7)</td>
<td>2</td>
</tr>
<tr>
<td>Readmission</td>
<td>18 (8.0)</td>
<td>1</td>
</tr>
<tr>
<td>Wound infection</td>
<td>11 (4.9)</td>
<td>4</td>
</tr>
<tr>
<td>Missed injury</td>
<td>14 (6.3)</td>
<td>1</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>11 (4.9)</td>
<td>1</td>
</tr>
<tr>
<td>Sepsis</td>
<td>8 (3.6)</td>
<td>1</td>
</tr>
<tr>
<td>Reintubation or unplanned intubation</td>
<td>13 (5.8)</td>
<td>0</td>
</tr>
<tr>
<td>Triage issue</td>
<td>6 (2.6)</td>
<td>0</td>
</tr>
<tr>
<td>Fall</td>
<td>8 (3.5)</td>
<td>0</td>
</tr>
<tr>
<td>Dislodged tube</td>
<td>6 (2.6)</td>
<td>0</td>
</tr>
<tr>
<td>Unplanned intensive care unit admission</td>
<td>6 (2.6)</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>44 (19.6)</td>
<td>22</td>
</tr>
<tr>
<td>Deaths</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths (including discharge to hospice)</td>
<td>84 (37.5)</td>
<td>_a</td>
</tr>
<tr>
<td>Unanticipated mortality with opportunity for improvement</td>
<td>12 (14.8)</td>
<td></td>
</tr>
<tr>
<td>Anticipated mortality with opportunity for improvement</td>
<td>17 (20.9)</td>
<td></td>
</tr>
<tr>
<td>Anticipated mortality without opportunity for improvement</td>
<td>52 (64.1)</td>
<td></td>
</tr>
<tr>
<td>Failure to rescue, n (%)</td>
<td>31 (13.8)</td>
<td>N/A b</td>
</tr>
</tbody>
</table>

_a Already mentioned.

_b Not applicable.

https://humanfactors.jmir.org/2019/3/e14819
Factors Associated with Complications

The most common factor associated with a complication was nature of injury (92/224 patients, 41.1%), followed by delays in intervention (41/224 patients, 18.3%). Moreover, 86 (86/224, 38.4%) patients with complications were described to have a mild level of harm associated with their complication, and 122 (122/224, 54.4%) patients had only 1 factor associated with a given complication. The median level of harm associated with a given complication was 1 (intraquartile range [IQR], 0-4; Table 2).

Table 2. Factors contributing to complications and level of harm in performance improvement patients (N=224).

<table>
<thead>
<tr>
<th>Factors contributing to complications</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of factor, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Delay in diagnosis</td>
<td>35 (15.6)</td>
</tr>
<tr>
<td>Delay in intervention</td>
<td>41 (18.3)</td>
</tr>
<tr>
<td>Error in diagnosis</td>
<td>5 (2.2)</td>
</tr>
<tr>
<td>Error in technique</td>
<td>24 (10.7)</td>
</tr>
<tr>
<td>Error in judgment</td>
<td>22 (9.8)</td>
</tr>
<tr>
<td>Patient refusal</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Nature of injury</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td><strong>Factor severity, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No harm</td>
<td>70 (31.3)</td>
</tr>
<tr>
<td>Mild harm</td>
<td>86 (38.4)</td>
</tr>
<tr>
<td>Moderate harm</td>
<td>42 (18.8)</td>
</tr>
<tr>
<td>Severe harm</td>
<td>12 (5.4)</td>
</tr>
<tr>
<td>Death</td>
<td>14 (6.3)</td>
</tr>
<tr>
<td><strong>Level of harm, median (intraquartile range)</strong></td>
<td>1 (0-4)</td>
</tr>
<tr>
<td><strong>Number of factors per complication, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>122 (54.5)</td>
</tr>
<tr>
<td>2</td>
<td>32 (14.3)</td>
</tr>
<tr>
<td>3</td>
<td>8 (3.6)</td>
</tr>
<tr>
<td>4</td>
<td>2 (0.9)</td>
</tr>
</tbody>
</table>

Regression Models

In our logistic regression model for level of harm, physician-attributed errors (odds ratio [OR] 2.82; \( P=.001 \), most commonly errors in technique, were associated with higher levels of harm, whereas delays in diagnosis or intervention were not significant in this analysis. Nature of injury was associated with higher level of harm (OR 1.91; \( P=.01 \)), whereas need for surgery was associated with decreased level of harm (OR 0.53; \( P=.02 \)). Patients with higher ISS on admission (OR 1.04; \( P<.001 \)) were more likely to have FTR. Each additional factor involved increased level of harm (OR 2.09; \( P<.001 \)) and nearly doubled likelihood of FTR (OR 1.95; \( P=.01 \)). Lipsitz goodness-of-fit test for the level of harm model demonstrated a \( P \) value of .26 (\( \chi^2 = 49.5 \)). For the FTR model, AOC SE was 0.048 (95% CI 0.71-0.9; \( P<.001 \)), and Hosmer-Lemeshow chi-square value was 11.3 with a degree of freedom of 9 (\( P=.18 \); Table 3).
Table 3. Logistic regression model for level of harm and failure to rescue.

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Level of harm</th>
<th>Failure to rescue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR^a (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Age</td>
<td>1.0 (0.9-1.0)</td>
<td>.59</td>
</tr>
<tr>
<td>Male</td>
<td>1.0 (0.6-1.7)</td>
<td>.99</td>
</tr>
<tr>
<td>Charlson score</td>
<td>1.0 (0.8-1.1)</td>
<td>.68</td>
</tr>
<tr>
<td>Delay</td>
<td>1.4 (0.8-2.6)</td>
<td>.20</td>
</tr>
<tr>
<td>Error</td>
<td>2.8 (1.5-5.2)</td>
<td>.001</td>
</tr>
<tr>
<td>Nature of injury</td>
<td>1.9 (1.1-3.3)</td>
<td>.01</td>
</tr>
<tr>
<td>Number of factors</td>
<td>2.1 (1.5-2.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Any surgery</td>
<td>0.5 (0.3-0.9)</td>
<td>.02</td>
</tr>
<tr>
<td>Injury severity score</td>
<td>1.0 (0.96-0.99)</td>
<td>.01</td>
</tr>
</tbody>
</table>

^aOR: odds ratio.

**Patient Characteristics**

Patients presented at PI meetings were more likely to be older (mean age 51.3 years, SE 1.58, vs 46.4 years, SE 0.51; P=.002). In addition, this group was more likely to present with lower GCS (median 14, IQR 3-15, vs median 15, IQR 15-15; P<.001) and higher ISS (median 22, IQR 11.75-29, vs median 8, IQR 4-10; P<.001). A higher proportion of patients presented at PI meetings were hypotensive on admission (7.1% vs 0.8%; P<.001). Finally, patients presented at PI meetings had a greater dispersion in Revised Trauma Score than patients without complications (median 5.64, IQR 1.02-5.64, mean 4.16, vs median 5.64, IQR 5.64-5.64, mean 5.5; P<.001). Trauma ISS was also lower in patients with complications (median 0.7, IQR 0.15-0.9, mean 0.56, vs median 0.94, IQR 0.84-0.97, mean 0.89; P<.001; Table 4).

Of those patients presented at PI meetings, the most common mechanism of injury was falls (102 patients, 102/224, 45.5%). Moreover, 97 (97/224, 43.3%) patients in this group were admitted to a step-down unit from the trauma bay, and 122 patients (122/224, 54.4%) underwent a procedure during their admission (Table 5).
Table 4. Patients’ demographics and physiological characteristics (patients presented at performance improvement versus other trauma patients).

<table>
<thead>
<tr>
<th>Patient demographics</th>
<th>All patients (N=2216)</th>
<th>Patients with complications (n=224)</th>
<th>Patients without complications (n=1992)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All age, mean (SE)</td>
<td>49.6 (22.9)</td>
<td>51.3 (1.58)</td>
<td>46.5 (0.51)</td>
<td>.01</td>
</tr>
<tr>
<td>Elderly patients (&gt;65), n (%)</td>
<td>484 (21.84)</td>
<td>61 (27.2)</td>
<td>423 (21.23)</td>
<td>.04</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1594 (71.93)</td>
<td>158 (70.5)</td>
<td>1436 (72.08)</td>
<td>.62</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>168 (7.58)</td>
<td>26 (11.6)</td>
<td>142 (7.12)</td>
<td>.02</td>
</tr>
<tr>
<td>Black</td>
<td>380 (17.14)</td>
<td>24 (10.7)</td>
<td>356 (17.87)</td>
<td>.01</td>
</tr>
<tr>
<td>Other</td>
<td>864 (38.98)</td>
<td>84 (37.5)</td>
<td>780 (39.15)</td>
<td>.63</td>
</tr>
<tr>
<td>Unknown</td>
<td>4 (0.18)</td>
<td>0 (0.0)</td>
<td>4 (0.20)</td>
<td>.50</td>
</tr>
<tr>
<td>White</td>
<td>800 (36.10)</td>
<td>90 (40.2)</td>
<td>710 (35.64)</td>
<td>.18</td>
</tr>
<tr>
<td>Insurance status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private, n (%)</td>
<td>946 (42.68)</td>
<td>89 (39.7)</td>
<td>857 (43.02)</td>
<td>.34</td>
</tr>
<tr>
<td>Public, n (%)</td>
<td>881 (39.75)</td>
<td>90 (40.2)</td>
<td>791 (39.70)</td>
<td>.89</td>
</tr>
<tr>
<td>Self-pay, n (%)</td>
<td>389 (17.55)</td>
<td>45 (20.1)</td>
<td>344 (17.26)</td>
<td>.29</td>
</tr>
<tr>
<td>Charlson score, median (IQR)</td>
<td>1 (0-2)</td>
<td>1 (0-4)</td>
<td>1 (0-2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Intensive care unit, length of stay (days), median (IQR)</td>
<td>0 (0-0)</td>
<td>1.77 (0-5.9)</td>
<td>0 (0-0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Vent days, median (IQR)</td>
<td>0 (0-0)</td>
<td>0 (0-2)</td>
<td>0 (0-0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Revised trauma score, median (IQR)</td>
<td>5.6 (5.6-5.6)</td>
<td>5.6 (1.0-5.6)</td>
<td>5.6 (5.6-5.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Trauma and injury severity score (N=2169), median (IQR)</td>
<td>0.9 (0.8-0.9)</td>
<td>0.7 (0.1-0.9)</td>
<td>0.9 (0.8-0.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Physiologic characteristics (n=170)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP&lt;sup&gt;b&lt;/sup&gt; (mm Hg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All patients, median (IQR)</td>
<td>136 (112-155.2)</td>
<td>134.5 (115.2-155)</td>
<td>136 (122-152)</td>
<td>.18</td>
</tr>
<tr>
<td>Hypotensive (SBP &lt;90), n (%)</td>
<td>31 (1.46)</td>
<td>16 (7.1)</td>
<td>15 (0.75)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>GCS&lt;sup&gt;c&lt;/sup&gt; (N=122)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All patients, median (IQR)</td>
<td>15 (15-15), 14.1</td>
<td>14 (3-15), 10.67</td>
<td>15 (15-15), 14.48</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>GCS&lt;8, n (%)</td>
<td>124 (5.86)</td>
<td>70 (31.2)</td>
<td>54 (2.71)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ISS&lt;sup&gt;d&lt;/sup&gt; (N=1947)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All patients, median (IQR)</td>
<td>8 (4-12)</td>
<td>22 (11.7-29)</td>
<td>8 (4-10)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ISS&gt;15, n (%)</td>
<td>376 (16.96)</td>
<td>150 (66.9)</td>
<td>226 (11.34)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>IQR: intraquartile range.

<sup>b</sup>SBP: systolic blood pressure.

<sup>c</sup>GCS: Glasgow Coma Scale.

<sup>d</sup>ISS: injury severity score.
Table 5. Injuries and procedures for patients presented at performance improvement (N=224).

<table>
<thead>
<tr>
<th>Injuries and procedures</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abbreviated injury score, median (intraquartile range)</strong></td>
<td></td>
</tr>
<tr>
<td>Head</td>
<td>3 (0-4)</td>
</tr>
<tr>
<td>Chest</td>
<td>0 (0-3)</td>
</tr>
<tr>
<td>Abdomen</td>
<td>0 (0-1.7)</td>
</tr>
<tr>
<td>Extremity</td>
<td>0 (0-2)</td>
</tr>
<tr>
<td><strong>Mechanism of injury, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Stab</td>
<td>8 (3.6)</td>
</tr>
<tr>
<td>Gunshot wound</td>
<td>11 (4.9)</td>
</tr>
<tr>
<td>Assault</td>
<td>8 (3.6)</td>
</tr>
<tr>
<td>Motor vehicle collision</td>
<td>9 (4.0)</td>
</tr>
<tr>
<td>Bicycle</td>
<td>14 (6.3)</td>
</tr>
<tr>
<td>Fall</td>
<td>102 (45.5)</td>
</tr>
<tr>
<td>Pedestrian struck</td>
<td>44 (19.6)</td>
</tr>
<tr>
<td>Motorcycle</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Others</td>
<td>26 (11.6)</td>
</tr>
<tr>
<td><strong>Disposition from trauma bay, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Step-down unit</td>
<td>97 (43.3)</td>
</tr>
<tr>
<td>Operating room</td>
<td>35 (15.6)</td>
</tr>
<tr>
<td>Intensive care unit</td>
<td>54 (24.1)</td>
</tr>
<tr>
<td>Monitored bed</td>
<td>24 (10.7)</td>
</tr>
<tr>
<td>IR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td>Floor</td>
<td>12 (5.4)</td>
</tr>
<tr>
<td><strong>Procedure, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Craniotomy</td>
<td>11 (4.9)</td>
</tr>
<tr>
<td>Open reduction and internal fixation</td>
<td>29 (12.9)</td>
</tr>
<tr>
<td>Exploratory laparotomy</td>
<td>20 (8.9)</td>
</tr>
<tr>
<td>None</td>
<td>102 (45.5)</td>
</tr>
<tr>
<td>IR</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Amputation</td>
<td>6 (2.7)</td>
</tr>
<tr>
<td>Thoracotomy</td>
<td>6 (2.7)</td>
</tr>
<tr>
<td>Intracranial pressure monitor</td>
<td>7 (3.1)</td>
</tr>
<tr>
<td>Chest tube insertion</td>
<td>14 (6.2)</td>
</tr>
<tr>
<td>Vascular surgery</td>
<td>4 (1.7)</td>
</tr>
<tr>
<td>Debridement or washout</td>
<td>7 (3.2)</td>
</tr>
<tr>
<td>Other</td>
<td>16 (21.4)</td>
</tr>
</tbody>
</table>

<sup>a</sup>IR: interventional radiology.

**Discussion**

**Principal Findings**

Care of trauma patients represents an environment that is prone to error. This is because of inherent illness of patients, time-sensitive decision making, and extensive handoffs and interplay of multiple specialties providing patient care. This was evident in this study as 51 errors contributed to a 8.93% complication rate (198 of 2216 patients) and 3.79% overall mortality rate (84 of 2216 patients) during the study period. Our reported mortality is in line with the 2% to 29% mortality that has been documented previously in the trauma literature [16,17]. Although reporting of complications is invariably center

https://humanfactors.jmir.org/2019/3/e14819

JMIR Hum Factors 2019 | vol. 6 | iss. 3 | e14819 | p.41

(page number not for citation purposes)
dependent because of the need for self-tracking and lack of consistent definitions, the rate of complications seen during our study period is also consistent with what has been reported [18,19]. To our knowledge, this is the first study in the trauma literature that focuses on factors that contribute to harm using the AHRQ system. In this study, increasing number of factors involved was significantly associated with increasing levels of patient harm. This is unsurprising, as the Swiss cheese model for error has shown that it is often multiple errors, not 1 single factor, that lead to harm for a given complication [20]. As trauma involves a complex interaction of clinicians and systems providing care, human factors inevitably affect the course of a critically injured patient. However, not all factors are created equal; in this study, certain errors were more likely to cause harm than others. Delay in diagnosis or intervention, for example, was not associated with a statistically increased level of harm or increased likelihood of FTR, whereas physician errors, either in diagnosis, technique, or judgment, were associated with increased harm (adjusted OR [AOR] 2.82, CI 1.52-5.25; P<.001). Similar provider errors have been described in preventable and potentially preventable deaths in the early resuscitation period and surgical intensive care unit, particularly when managing unstable patients, hemorrhagic shock, and threatened airways at major US trauma centers [10-12].

Several studies in the literature have previously linked patient-related factors with complications after traumatic injury. Bell et al demonstrated that preexisting comorbidities contributed significantly to mortality after complication in a trauma population [21], whereas others have shown that insurance status [22] and age [23] are associated with increased likelihood of FTR. In our analysis, we demonstrated that physician-related factors are more strongly associated with an increased risk for harm compared with underlying patient and injury attributes. We also demonstrate that error compounding significantly contributes to harm after complications with each increase in number of factors, effectively doubling the level of harm (AOR 2.09, 95% CI 1.52-2.91).

FTR represents a hospital’s inability to rescue a patient from complications. The FTR rate in our patient population was 18.8% (31/224). This is consistent with the literature, which cites FTR rates ranging from 6.8% to 19.8% [6,24,25]. Studies have examined specific factors that contribute to FTR and patient harm. Joseph et al [24] found that patients’ age, trauma mechanism, insurance status, and number of blood products administered on the second day of hospitalization significantly contributed to likelihood of FTR. Bell et al [21] found that uninsured patients had the lowest likelihood of developing a complication (OR 0.86), and yet they were more likely to experience FTR (OR 1.34) than patients who were privately insured (OR 1.25) or publicly insured (OR 1.17). Another study demonstrated that hospital- and physician-based factors, such as anesthesia board certification and presence of surgical house staff, were associated with FTR, whereas severity of illness was not [6]. Our analysis only demonstrated 2 significant factors for FTR in our regression model, although this may have been limited by sample size. Increased number of errors was associated with an almost 2-fold increase in FTR (AOR 1.95, 95% CI 1.16-3.27). ISS was the only patient-intrinsic factor identified and associated with a 4% increase per point increase in ISS (AOR 1.04, 95% CI 1.02-1.07).

Other findings, which are counter intuitive, are the inverse association between increasing ISS and need for immediate operative intervention and level of harm. One would infer that patients who required a surgery for their injuries had more severe injuries or were more critically ill, yet this population had a lower associated level of harm from a given complication. This suggests that, perhaps, there is a higher level of vigilance in this group of patients compared with those less injured or that physician-attributable errors significantly increased odds of increased level of harm more so than patient-based nature of injury. This suggests that there may be more opportunities for overall harm prevention.

Limitations

There were several limitations present in this study. As this study was retrospective in nature, it was not designed to prove causality between the relationships and associations between harm and outcomes demonstrated. We are only able to make generalizations based on the experiences of a single trauma center. Reporting of complications is voluntary and lacks uniformity; however, we believe that we are fortunate to have a PI coordinator who is invaluable in assisting the trauma service in identifying as many complications and adverse events as possible. The level of training of the individuals making the error or service responsible for the error was not captured; this highlights the difficulty in assigning attribution for postinjury complications. This study was also limited by a small sample size of patients, especially patients who were deemed FTR and may predispose to type II error. It would have been ideal to capture where in the course of evaluation or postinjury course that complications occurred (ie, resuscitation, operating room, or acute care), but this information was not available. In addition, physicians were asked to record factors contributing to harm, and these may be subjective measures, even if reached by consensus. Finally, it is possible that our analysis may be affected by hindsight bias. Surgeons at our institution may have been more likely to find more contributing factors or errors with a poor outcome such as a mortality or serious complication when reviewing the case in PI meetings. Unfortunately, this is a limitation of this retrospective case review.

Conclusions

This analysis applies the current concept of FTR and patient complication prevention to the trauma patient population. We have demonstrated that the increasing number of errors identified in patient care directly correlates with level of harm seen after traumatic injury. Interestingly, certain types of errors are more associated with harm; in particular, physician-attributable errors are more strongly correlated with harm than underlying patient factors.

This is, to our knowledge, one of the first studies to categorize level of harm using new AHRQ guidelines. Future studies should examine interventions that could prevent or mitigate physician-attributable errors. These could be further classified into error type (ie, skill-, rule-, or knowledge-based) to further assess which errors carry more weight or risk to patient harm.
This information could then be used to develop further provider training or trauma system enhancement for quality improvement. Prospective evaluation of these specific interventions could then be used to assess their impact on patient-related complications and levels of harm.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Performance improvement form.

References


Abbreviations

AHRQ: Agency for Healthcare Research and Quality
AOR: adjusted odds ratio
BHC: Bellevue Hospital Center
FTR: failure to rescue
GCS: Glasgow Coma Scale
IQR: intraquartile range
ISS: injury severity score
OFI: opportunity for improvement
PI: performance improvement
TOPIC: Trauma Outcomes and Performance Improvement Course

©Elizabeth Warnack, Hersch Leon Pachter, Beatrix Choi, Charles DiMaggio, Spiros Frangos, Michael Klein, Marko Bukur. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 26.09.2019 This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
Evaluating the Feasibility of a Software Prototype Supporting the Management of Multimorbid Seniors: Mixed Methods Study in General Practices

Christine Kersting¹, Dr rer medic; Birgitta Weltermann¹,², MD, MPH

¹Institute for General Medicine, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
²Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn, Bonn, Germany

Abstract

Background: Longitudinal, patient-centered care represents a challenge for general practices. Decision support and reminder systems can offer targeted support.

Objective: The objective of this study was to follow a user-oriented, stepwise approach to develop an add-on for German electronic health record (EHR) systems, which aims to support longitudinal care management of multimorbid seniors, using a flag system displaying patient-centered information relevant for comprehensive health care management. This study evaluated the prototype’s feasibility from both a technical and users’ perspective.

Methods: The study was conducted with 18 general practitioners (GPs) and practice assistants (PAs) from 9 general practices using a mixed methods approach. In all practices, 1 GP and 1 PA tested the software each for 4 multimorbid seniors selected from the practice patient data. Technical feasibility was evaluated by documenting all technical problems. To evaluate the feasibility from the users’ perspective, participants’ responses during the software test were documented. In addition, they completed a self-administered questionnaire, including the validated System Usability Scale (SUS). Data were merged by transforming qualitative data into quantitative data. Analyses were performed using univariate statistics in IBM SPSS statistics.

Results: From a technical perspective, the new software was easy to install and worked without problems. Difficulties during the installation occurred in practices lacking a 64-bit system or a current version of Microsoft .NET. As EHRs used in German practices do not provide an interface to extract the data needed, additional software was required. Incomplete flags for some laboratory data occurred, although this function was implemented in our software as shown in previous tests. From the users’ perspective, the new add-on provided a better overview of relevant patient information, reminded more comprehensively about upcoming examinations, and better supported guideline-based care when compared with their individual practice strategies. A total of 14 out of 18 participants (78%) were interested in using the software long-term. Furthermore, 8 of 9 GPs were willing to pay 5 to 25 Euros (mean 14.75, SD 5.93) monthly for its use. The usability was rated as 75% (43%-95%).

Conclusions: The new EHR add-on was well accepted and achieved a good usability rating measured by the validated SUS. In perspective, the legally consolidated, standardized interface to German EHRs will facilitate the technical integration. In view of the high feasibility, we plan to study the software’s effectiveness in everyday primary care.

Trial Registration: German Clinical Trials Register DRKS00008777; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00008777

(JMIR Hum Factors 2019;6(3):e12695) doi:10.2196/12695
Introduction

Background

Health care management of seniors (persons aged ≥ 65 years) is complex as more than 55% are multimorbid [1]. Electronic health records (EHRs) have the potential to support care management, but were shown to insufficiently support longitudinal patient-centered management as they (1) require high user-system interaction and work slow, (2) lack user-friendliness and orientation to typical health care processes, (3) provide insufficient interoperability, (4) offer insufficient service if help is needed, and (5) inadequately distinguish between information relevant and irrelevant for patient care [2-7]. Surveys evaluating the use and functional capacity of EHR in German primary care revealed a high implementation rate of EHR systems but little multifunctional capacity, especially with regard to reminder and recall systems [2,6,8,9].

Considering that the demographic change will lead to an increasing number of seniors and complex, multimorbid patients, health information technologies (IT), which adequately support health care management, are needed. EHR add-ons such as clinical decision support systems (CDSS), including reminder systems, can provide targeted support as they were shown to effectively support management of, for example, diabetes or hypertension [10-14]. Limiting the benefit of such disease-specific CDSS, a survey among primary care physicians revealed that physicians are more willing to accept new IT solutions, when supporting the management of elderly patients with multiple conditions or polypharmacy [15]. In addition, many systems do not adequately meet physicians’ needs, even though an early involvement of the target group into software development is described as key facilitator for software acceptance [16-18] and is therefore highly recommended [19].

Objectives

To overcome deficits of the current EHR solutions, we developed a new software add-on for German EHR systems. The add-on named eCare*Seniors aims to provide a quick and intuitive overview of each patient’s care needs to adequately support comprehensive and longitudinal patient-centered care management not only for multimorbid seniors in general practices, but for other patient groups as well [20,21]. To provide a quick overview of each patient’s care needs, relevant information is extracted from a patient’s EHR and reedited using so-called flags, which are a combination of colored fields and short keywords (Figure 1) [20,21]. The theoretical concept is based on the chronic care model [22-24], whereas from a functional perspective, eCare*Seniors is classified as a CDSS, integrating the functions of a reminder, advisor, and critic [25-27]. Further details on the new add-on and its functions are published elsewhere [20].

Figure 1. Example of the patient-centered flag system of eCare*Seniors displaying information relevant to health care management. Note: As the software was developed for German general practitioners’ practices, the screenshot is in German, but includes an English explanation.
Following the recommendation for the development of health IT [19], eCare*Seniors was developed using a user-oriented, stepwise approach. First, an assessment of the status quo showed that German general practitioners (GPs) and practice assistants (PAs) are not satisfied with current EHR solutions and had started to self-design reminder systems within the EHR to maintain an overview of each patient’s care [5,28]. In a second step, we presented the concept for the new software add-on to GPs and PAs who generally welcomed the approach but expressed a desire for configuration options to adapt the software to their individual practice needs. On the basis of this target group information, we then implemented the software prototype in a third step. In this paper, we present the results of the feasibility study of the new software add-on, which was the fourth step of the development: the software was tested in GP practices to assess the feasibility from both a technical and users’ perspective.

Methods

Study Design

The feasibility study was conducted using a simultaneous mixed methods design involving qualitative and quantitative approaches [29]. According to recommendations on how to design feasibility studies, we focused on the 5 key criteria, integration, implementation, acceptability, demand, and practicability [30], which we subdivided into 2 categories:

1. Technical feasibility: Integration of the new add-on into existing IT infrastructure and implementation of the new add-on with regard to its success or failure of execution during the work process.
2. Feasibility from the users’ perspective: Demand for and acceptance of the new add-on within the target group, and assessment of the usability from the users’ perspective. Usability as one of the most important indicators for feasibility was measured using the validated System Usability Scale (SUS) [31,32].

Setting and Recruitment

As recommendations on how to perform usability studies recommend sample sizes of 3 to 20 participants and the SUS was even shown to provide consistent results in study groups of 12 persons [33,34], we targeted a sample size ≥12 participants. The study was performed in a convenience sample of GP practices of the practice networks of the Institute for General Medicine, University of Duisburg-Essen, Germany, and the Institute for Family Medicine and General Practice of the University of Bonn, Germany (both >100 practices), with 1 GP and 1 PA each. Owing to the lack of a standardized, open interface of existing EHRs, a second software was used to extract the data needed for our new software add-on. To assure this connectivity to existing EHRs, only practices that apply one of 13 specific EHR solutions, which are used by about 85% of all German GP practices [35], were eligible to participate. As the new software is meant to be usable and feasible for every practice, no further inclusion and exclusion criteria were defined. Considering a participation rate of at least 40%, 30 participants from 15 practices who met the inclusion criteria and were located in the Greater Essen region, Germany, or those who already participated in an earlier step of the study were contacted and invited to participate.

Data Collection and Data Management

All participating GP practices were visited twice by a research team member of the institute. During the first visit, the researcher installed the prototype of the newly developed add-on eCare*Seniors. During the second visit, the new add-on was introduced to the practice and tested by 1 GP and 1 PA per practice with patient data. To evaluate the technical feasibility, all problems that occurred during the installation process and when executing the new add-on during the test were documented. To evaluate the feasibility from the users’ perspective, data were collected using a prestructured interview guide for documentation of the software test and a self-administered questionnaire, which was completed by each participant after the software test:

1. Documentation of the software test: First, the GP was asked to configure the software according to the practice’s needs. The preferences were documented using a checklist. Afterward, the software was tested independently by the GP and the PA. Both accessed 4 complex patients (defined as seniors aged 65 years or above with 2 or more chronic conditions) within the prototype of the new add-on and within the practice’s EHR and compared both systems based on the following questions:
   - Is there any information available at first glance in the new add-on? (yes/no)
   - Does the new add-on display any information at first glance which you do not see instantly in the patient’s EHR? (yes/no)
     - What information? (open answer)
     - Is this information important? (yes/no)
   - Is there any information available at first glance in the patient’s EHR which you do not see instantly in the new add-on? (yes/no)
     - What information? (open answer)
     - Is this information important? (yes/no)

   To complete the test, GPs and PAs clicked through the remaining menu of the software add-on and applied the method of thinking aloud to help the researcher identify problems regarding comprehensibility and handling.

2. Written questionnaire including the following aspects:
   - German version of the validated SUS to evaluate the usability of software solutions based on 10 questions on a 5-point Likert scale [31,32]
   - Two open-ended questions to request which aspects of the software were liked most and which aspects needed to be optimized
   - One closed-ended question to assess whether participants would choose to use the new add-on long-term
   - Five questions comparing the new software add-on with the current practice-specific care management strategies to assess the aspects of quick overview, reminder functionality, support of individual, guideline-oriented care planning, support of time-efficient processes, and financial benefits on a 5-point Likert scale

https://humanfactors.jmir.org/2019/3/e12695/
One open-ended question to assess how much money (in Euros) the potential users would be willing to pay for such software per month.

In addition, sociodemographic characteristics of the participants were assessed.

All data were entered manually in an access-restricted database in IBM SPSS Statistics for Windows, version 24 (IBM Corp).

**Data Analysis**

Following an approach of Creswell and Plano Clark for integrating quantitative and qualitative data during the analysis process, we merged the data by transforming qualitative into quantitative data using categorization [29]. After that, all data were analyzed using descriptive statistics in SPSS. The usability was analyzed by determining the mean SUS score. The score can assume values between 0 and 40. Results were interpreted in percentage, that is, each score was multiplied by the factor 2.5 [31,32]. A SUS score ≥70% denotes good usability, whereas a score ≤50% indicates a considerable need for improvement [32].

**Ethical Considerations**

Ethical approval was obtained from the Ethics Committee of the Medical Faculty of the University of Duisburg-Essen (reference number: 14-5980-BO; date of approval: June 1, 2015). All participants received written information and signed the informed consent forms, which are stored at the institute. All members of the practice network had provided written informed consent for the pseudonymized analysis of data on practice characteristics within the scope of scientific research of the institute.

**Results**

**Study Characteristics**

Of the 30 participants from 15 practices invited, 6 participants from 3 practices refused to participate, yielding an initial response rate of 80% (24/30). As detailed below, 3 practices with 6 participants willing to participate used an operating system, which was not supported by the new software (32-bit instead of the required 64-bit). Finally, 18 participants (9 GPs, 9 PAs) from 9 practices took part. The practices used 4 different EHR software solutions covering about 26% of all EHR solutions installed in German GP practices [35]. All practices participated with 1 GP and 1 PA each (N=18). Furthermore, 5 of these practices were group practices. The GPs were aged between 40 and 56 years (mean 47.4, SD 5.2); 8 of them were male. All PAs were female and aged between 22 and 60 years (mean 36.7, SD 14.3). All participants tested the software add-on with 4 patient charts, yielding a total of 72 patient tests.

**Technical Feasibility**

The following technical problems were observed during the installation processes:

- As mentioned above, the new software could not be installed in 3 practices as they used a 32-bit operating system rather than the required 64-bit system.
- Installation was delayed in 4 practices as their Microsoft .NET program, which is required to run the desktop application of our software, needed to be updated.
- One of the practices did not have sufficient administration rights to install the software. To continue the installation process, access data had to be requested from the responsible software publisher.

Once these installation problems were resolved, the software prototype worked without problems. No technical problems occurred when executing the software during the practice tests. Nevertheless, in all 9 practices, few flags were not displayed in eCare*Seniors. In all 9 practices, the present year was named as due date in all flags on upcoming vaccinations. Compared with the information available within the patient charts, flags on chronic medication were incomplete in 4 practices, those on chronic diseases in 3 practices, and those on laboratory values in 1 practice. As all these flags were completely and successfully implemented in tests conducted during the software development, these problems are likely because of connectivity problems of the second software used to extract EHR data.

**Feasibility From the Users’ Perspective**

During the practice test, each participant compared the information displayed at first glance in the practice’s EHR with the flags presented by eCare*Seniors. For 69 of the 72 (96%) patient charts selected for the test, participants stated that information which they consider important for comprehensive care management was available at first glance only in eCare*Seniors but not in their EHR (Table 1).

During the interviews alongside the chart tests, for 54 of the 72 (75%) patient cases tested, participants named some additional information which they would like to be presented in eCare*Seniors. The researchers judged the following information as reasonable to be evaluated for integration into a future version of eCare*Seniors:

- More detailed information on chronic diseases, for example, renal impairment in diabetic patients (mentioned in 14 of the 54 patient charts, 26%)
- More space for freely formulated notes and risks (12/54, 22%)
- More detailed information on allergies (11/54, 20%)
- More detailed information on chronic medication (9/54, 17%)
- Residential care/home visit patient (8/54, 15%)
- More detailed information on physical parameters such as height and weight (5/54, 9%)

Other wishes mentioned need be evaluated with regard to technical implementability (eg, the availability of an electronic medication plan or the date of the next disease management program examination).

Furthermore, 8 GPs and 8 PAs each (16/18, 89%) specified one or more aspect of the new software they liked most (Figure 2) and 9 GPs and 7 PAs (16/18, 89%) specified one or more aspect which needed to be optimized (Figure 3).
### Table 1. Practice test: information only available at first glance in the new software and its relevance for care management.

<table>
<thead>
<tr>
<th>Information displayed in eCare*Seniors</th>
<th>Information was not available in patients’ electronic health records at first glance (N=69), n (%)</th>
<th>Information was judged relevant for care management of this patient n (%)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upcoming vaccinations</td>
<td>64 (93)</td>
<td>60 (94)</td>
<td>64</td>
</tr>
<tr>
<td>Upcoming cancer screening including referrals to specialists</td>
<td>33 (48)</td>
<td>28 (85)</td>
<td>33</td>
</tr>
<tr>
<td>Laboratory findings</td>
<td>12 (17)</td>
<td>12 (100)</td>
<td>12</td>
</tr>
<tr>
<td>Chronic diseases (eg, rheumatoid arthritis)</td>
<td>8 (12)</td>
<td>8 (100)</td>
<td>8</td>
</tr>
<tr>
<td>Warnings/allergies</td>
<td>4 (6)</td>
<td>4 (100)</td>
<td>4</td>
</tr>
<tr>
<td>Billing codes</td>
<td>3 (4)</td>
<td>3 (100)</td>
<td>3</td>
</tr>
<tr>
<td>Geriatric care complex</td>
<td>2 (3)</td>
<td>2 (100)</td>
<td>2</td>
</tr>
<tr>
<td>Participation in a disease management program</td>
<td>1 (1)</td>
<td>1 (100)</td>
<td>1</td>
</tr>
</tbody>
</table>

### Figure 2. Positive aspects of eCare*Seniors according to general practitioners and practice assistants.

<table>
<thead>
<tr>
<th>General practitioners (n=8)</th>
<th>Practice assistants (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well structured, clear and not too complex</td>
<td></td>
</tr>
<tr>
<td>Need-based configuration options</td>
<td></td>
</tr>
<tr>
<td>All relevant information at a glance</td>
<td></td>
</tr>
<tr>
<td>Good overview of vaccinations</td>
<td></td>
</tr>
<tr>
<td>Good reminder about preventive measures</td>
<td></td>
</tr>
<tr>
<td>Visualization with colours</td>
<td></td>
</tr>
<tr>
<td>Simple handling</td>
<td></td>
</tr>
<tr>
<td>Helps save time</td>
<td></td>
</tr>
<tr>
<td>Stressing of laboratory values</td>
<td></td>
</tr>
<tr>
<td>Good idea or good approach</td>
<td></td>
</tr>
</tbody>
</table>
The results of the questionnaire survey showed that 14 of the 18 participants (78%; 7 GPs and 7 PAs) wished to use the new software long-term. In addition, 3 GPs even stated that they would implement and use the software immediately, if it was linked directly to the EHR software and did not require a second software for bridging the interface. Despite its design to support care management of seniors, the configuration options of eCare*Seniors allowed for applying the software for other patient groups as well. This option was selected by 7 of the 9 practices; 1 practice extended the use to patients aged ≥35 years, 6 to all patients. eCare*Seniors was rated positively in comparison with existing patient management strategies of the individual practices (Table 2). The 2 GPs and 2 PAs who did not express a desire to use the new software in the long term said they were already using a well-functioning system which offered the same functionality. All GPs except one were willing to pay a monthly license fee for the software, ranging from 5 to 25 Euros (mean 14.75, SD 5.93).

Table 2. Questionnaire answers of general practitioners and practice assistants: Comparison of eCare*Seniors with the current practice-specific patient management approach (N=18). Percentages are reported for valid cases.

<table>
<thead>
<tr>
<th>Aspects assessed</th>
<th>Strongly disagree, n (%)</th>
<th>Disagree, n (%)</th>
<th>Neither agree nor disagree, n (%)</th>
<th>Agree, n (%)</th>
<th>Strongly agree, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eCare*Seniors offers a quicker overview of all important patient-centered contents of care</td>
<td>1 (6)</td>
<td>4 (22)</td>
<td>1 (6)</td>
<td>9 (50)</td>
<td>3 (17)</td>
</tr>
<tr>
<td>I think that eCare*Seniors gives better reminders about upcoming checkups, vaccinations, and routine examinations</td>
<td>0 (0)</td>
<td>3 (17)</td>
<td>1 (6)</td>
<td>5 (28)</td>
<td>9 (50)</td>
</tr>
<tr>
<td>I think that eCare*Seniors better supports individual, guideline-oriented care planning for complex patients</td>
<td>0 (0)</td>
<td>3 (17)</td>
<td>2 (11)</td>
<td>7 (39)</td>
<td>6 (33)</td>
</tr>
<tr>
<td>I think that eCare*Seniors better supports time-efficient processes in time-limited everyday practice</td>
<td>0 (0)</td>
<td>2 (11)</td>
<td>5 (28)</td>
<td>3 (17)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>I think that the reminder function of eCare*Seniors offers a financial benefit</td>
<td>1 (6)</td>
<td>3 (17)</td>
<td>2 (11)</td>
<td>9 (50)</td>
<td>3 (17)</td>
</tr>
</tbody>
</table>
Table 3. Participants’ evaluation of the practicability based on the System Usability Scale (N=18). Percentages are reported for valid cases.

<table>
<thead>
<tr>
<th>Items of the System Usability Scale</th>
<th>Strongly disagree, n (%)</th>
<th>Disagree, n (%)</th>
<th>Neither agree nor disagree, n (%)</th>
<th>Agree, n (%)</th>
<th>Strongly agree, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think that I would like to use this system frequently</td>
<td>0 (0)</td>
<td>4 (22)</td>
<td>3 (17)</td>
<td>6 (33)</td>
<td>5 (28)</td>
</tr>
<tr>
<td>I found the system unnecessarily complex</td>
<td>10 (56)</td>
<td>3 (17)</td>
<td>3 (17)</td>
<td>2 (11)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>I thought the system was easy to use</td>
<td>0 (0)</td>
<td>1 (56)</td>
<td>2 (11)</td>
<td>7 (39)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>I think that I would need the support of a technical person to be able to use this system</td>
<td>7 (39)</td>
<td>8 (44)</td>
<td>1 (6)</td>
<td>0 (0)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>I found the various functions in this system were well integrated</td>
<td>0 (0)</td>
<td>3 (17)</td>
<td>4 (22)</td>
<td>10 (56)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>I thought there was too much inconsistency in this system</td>
<td>5 (28)</td>
<td>3 (17)</td>
<td>3 (17)</td>
<td>7 (39)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>I would imagine that most people would learn to use this system very quickly</td>
<td>0 (0)</td>
<td>1 (6)</td>
<td>0 (0)</td>
<td>10 (56)</td>
<td>7 (39)</td>
</tr>
<tr>
<td>I found the system very cumbersome to use</td>
<td>12 (67)</td>
<td>4 (22)</td>
<td>0 (0)</td>
<td>2 (11)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>I felt very confident using the system</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>3 (17)</td>
<td>7 (39)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>I needed to learn a lot of things before I could get going with this system</td>
<td>11 (65)</td>
<td>3 (18)</td>
<td>2 (12)</td>
<td>1 (6)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

The average SUS score was 75% (SD 14%), varying from 43% to 95%. A total of 14 of the 18 participants (78%) rated the usability as ≥70%, which denotes good usability, although 1 participant (5%) rated the usability as <50%. Stratified for GPs and PAs, the respective SUS score was 78% (SD 16%; range 43%-95%) and 73% (SD 12%; range 53%-93%). Evaluation results on item level are illustrated in Table 3.

Discussion

Principal Findings

From the users’ perspective, the new EHR add-on meets its intended purpose; GPs and PAs positively emphasize the intuitive, quick, and comprehensive overview of relevant, patient-centered information provided by eCare*Seniors. Applying the feasibility criteria defined by Bowen et al [30], the new add-on is generally feasible; it meets the users’ demands, is accepted within the target group, has a good usability rating, and no technical problems are encountered during the implementation, as long as practices use a 64-bit operating system. Currently, the biggest challenge is the connectivity between the new add-on and the existing IT infrastructure. In perspective, this will be facilitated by standardized and open interfaces to EHR solutions, which are legally consolidated in Germany since 2017 [36].

In the standardized and validated SUS scale, users rated the usability of eCare*Seniors as 75%, which denotes good usability [32]. As no similar IT approaches were studied in German general practices, the usability cannot be compared with that of other IT solutions. In the literature, only few studies report on user-oriented usability evaluations of newly developed electronic tools supporting patient-centered care management. To a minor extent, comparable approaches are described in 4 studies from the United States and Canada. Furthermore, 3 of these studies assessed the usability of newly developed, EHR-integrated CDSS supporting the standardized, guideline-oriented therapy of patients with chronic pain [37,38] or arterial fibrillation [39] in samples of 4 to 12 GPs and PAs in real primary care scenarios; other than eCare*Seniors these systems only support disease-specific patient management rather than comprehensive, patient-centered management [37-39]. Similar to eCare*Seniors, an approach from the United States aims at priority-oriented restructuring of patient-centered information within the EHR to provide a quick overview of patient’s care needs. On the basis of a previous study on information needs, 4 options for information presentation were developed and tested by 16 physicians using fictive patient cases. However, to our knowledge, this concept was only realized as prototype and has not been transferred to EHR solutions thus far [40]. Despite the limited comparability, these studies report similar usability ratings based on the SUS scores or comparable items [37-40].

Although eCare*Seniors predominantly aims to support health care management of multimorbid seniors, most practices wanted to use it for all their patients. This shows that the demand is not limited to complex patients only, which might be explained by the fact that the current EHR solutions were described as inconvenient, requiring high interaction, and lacking reminders relevant for care [5]. Interestingly, this result contradicts a study by Sittig et al indicating that physicians prefer IT solutions supporting the management of elderly patients with multiple conditions or polypharmacy [15]. Nevertheless, it has to be considered that our approach aimed to overcome deficits of German EHR systems. This also makes it difficult to compare our approach with other studies on CDSS as most studies evaluated the effect of newly developed, often imposed systems instead of further developing existing IT solutions with regard to deficits and users’ needs [10-14].

Limitations

The key strength of the development and feasibility testing of eCare*Seniors is the user-oriented, bottom-up approach, which facilitated adequate consideration and realization of the users’ demands. In addition, the mixed methods design involved qualitative and quantitative elements allowing for a more detailed assessment of the users’ opinions than a strict quantitative approach. For feasibility testing, the study sample
included 18 GPs/PAs and 4 patient cases per participant. Although this is within the range of sample sizes recommended for usability studies [33], the sample analyzed is too small to generate generalizable and reproducible results. Another aspect which limits the results' generalizability and reproducibility is the fact that the SUS provides self-reported, subjective measures of usability and the study lacks objective usability measures. In addition, the results might be limited because of a selection and response bias. First, it cannot be excluded that the participating practices had a greater interest in new health IT or had a higher affinity for technology than those who refused to participate. Second, the responses of the participants might be positively influenced by the presence of a member of the research team. As GPs and PAs expressed their criticism ad hoc, this bias is considered negligible.

Conclusions
This feasibility study shows that the newly developed EHR add-on is well accepted and usable; however, the technical integration into present IT infrastructures of general practices could be facilitated. To determine whether the utilization of the software positively influences practice organization, patient care, and patients' health outcomes, randomized controlled intervention studies are needed.

Acknowledgments
The authors thank all PAs and GPs for supporting this study. The authors also thank Marius Sharma, diploma computer scientist, for the technical realization of the software prototype and for his technical support during the feasibility study.

Conflicts of Interest
None declared.

References

https://humanfactors.jmir.org/2019/3/e12695/

https://humanfactors.jmir.org/2019/3/e12695/


Abbreviations

CDSS: clinical decision support system
EHR: electronic health record
GP: general practitioner
IT: information technology
PA: practice assistant
SUS: System Usability Scale
Understanding the Situated Roles of Electronic Medical Record Systems to Enable Redesign: Mixed Methods Study

Samar Helou¹, MSc; Victoria Abou-Khalil¹, MSc; Goshiro Yamamoto², PhD; Eiji Kondoh³, MD; Hiroshi Tamura⁴, MD; Shusuke Hiragi², MD, MBA; Osamu Sugiyama⁵, PhD; Kazuya Okamoto², PhD; Masayuki Nambu⁵, PhD; Tomohiro Kuroda², PhD

¹Department of Social Informatics, Graduate School of Informatics, Kyoto University, Kyoto, Japan
²Division of Medical Information Technology and Administration Planning, Kyoto University Hospital, Kyoto, Japan
³Department of Gynecology and Obstetrics, Graduate School of Medicine, Kyoto University, Kyoto, Japan
⁴Center for Innovative Research and Education in Data Science, Kyoto University, Kyoto, Japan
⁵Preemptive Medicine and Lifestyle Related Diseases Research Center, Kyoto University Hospital, Kyoto, Japan

Corresponding Author:
Samar Helou, MSc
Department of Social Informatics
Graduate School of Informatics
Kyoto University
Yoshida-honmachi, Sakyoku
Kyoto, 606-8501
Japan
Phone: 81 5055346505
Email: helou.samar@gmail.com

Abstract

Background: Redesigning electronic medical record (EMR) systems is needed to improve their usability and usefulness. Similar to other artifacts, EMR systems can evolve with time and exhibit situated roles. Situated roles refer to the ways in which a system is appropriated by its users, that is, the unintended ways the users engage with, relate to, and perceive the system in its context of use. These situated roles are usually unknown to the designers as they emerge and evolve as a response by the users to a contextual need or constraint. Understanding the system’s situated roles can expose the unarticulated needs of the users and enable redesign opportunities.

Objective: This study aimed to find EMR redesign opportunities by understanding the situated roles of EMR systems in prenatal care settings.

Methods: We conducted a field-based observational study at a Japanese prenatal care clinic. We observed 3 obstetricians and 6 midwives providing prenatal care to 37 pregnant women. We looked at how the EMR system is used during the checkups. We analyzed the observational data following a thematic analysis approach and identified the situated roles of the EMR system. Finally, we administered a survey to 5 obstetricians and 10 midwives to validate our results and understand the attitudes of the prenatal care staff regarding the situated roles of the EMR system.

Results: We identified 10 distinct situated roles that EMR systems play in prenatal care settings. Among them, 4 roles were regarded as favorable as most users wanted to experience them more frequently, and 4 roles were regarded as unfavorable as most users wanted to experience them less frequently; 2 ambivalent roles highlighted the providers’ reluctance to document sensitive psychosocial information in the EMR and their use of the EMR system as an accomplice to pause communication during the checkups. To improve the usability and usefulness of EMR systems, designers can amplify the favorable roles and minimize the unfavorable roles. Our results also showed that obstetricians and midwives may have different experiences, wants, and priorities regarding the use of the EMR system.

Conclusions: Currently, EMR systems are mainly viewed as tools that support the clinical workflow. Redesigning EMR systems is needed to amplify their roles as communication support tools. Our results provided multiple EMR redesign opportunities to improve the usability and usefulness of EMR systems in prenatal care. Designers can use the results to guide their EMR redesign activities and align them with the users’ wants and priorities. The biggest challenge is to redesign EMR systems in a way that amplifies their favorable roles for all the stakeholders concurrently.
**Introduction**

Enabling Redesign by Understanding the Situated Roles of an Electronic Medical Record System

The usability and usefulness of electronic medical record (EMR) systems are critical for their acceptance and effective use [1-3]. Accordingly, multiple user and usability studies were conducted with the aim of refining the systems’ functional and nonfunctional specifications [4-15]. Previous studies used interviews, surveys, focus groups, and observations to identify the needs of EMR users and the issues they encounter when using EMR systems. However, EMR users may have needs that they are not aware of or cannot articulate. Moreover, the experts’ verbal description of their work could be inconsistent with how they perform it in the field [16].

To address these limitations, we adopted a novel approach for finding EMR redesign opportunities. The approach is based on the idea of redesigning from appropriation [17]. Following this approach, we analyzed the EMR system as an artifact that evolves with time and exhibits situated roles. Situated roles refer to the ways in which a system is appropriated by its users, that is, the unintended ways the users engage with, relate to, and perceive the system in its context of use. The conceptual approach is depicted in Figure 1. Multiple EMR situated roles can exist; for example, one could use it as an explanation support tool to communicate information to a patient during a consultation or as an excuse to take a break from the conversation. These situated roles are usually unknown to the designers as they emerge with time as a response by the users to a contextual need or constraint. Understanding the situated roles of the system and the users’ attitudes regarding them could enable user-centered redesign opportunities.

We applied our approach in prenatal care settings, a unique health care setting that does not fit into the common clinician-patient scheme where EMR systems are usually studied. Prenatal care is the periodic care that a pregnant woman receives during her pregnancy. Unlike other health care settings where the aim is to address a patient’s health problems, the main goal of prenatal care is the prevention and early detection of diseases that can affect the pregnant woman and her fetus(es) [18]. Therefore, prenatal care is the setting in which we collect information about the health of an individual for the first time. The effective use of EMR systems in antenatal care is needed if we aim to have complete longitudinal health records.

**Purpose of This Study**

The purpose of this study was to identify EMR redesign opportunities in prenatal care settings. We achieved this purpose by answering the following research questions:

- What are the situated roles of EMR systems in prenatal care settings?
- Do users want to experience the situated roles more, or less, frequently?
- How important are the different situated roles to the users?

Using our results, designers can align their EMR redesign activities with the wants and priorities of EMR users in prenatal care.

**Prenatal Care in Japan**

In Japan, prenatal care is standardized by the Japan Society of Obstetrics and Gynecology and the Japan Association of Obstetricians and Gynecologists. Most of the pregnant women in Japan attend the recommended regular prenatal checkups. Women with uncomplicated pregnancies usually receive 14 checkups. The checkups start around their 8th week of pregnancy and finish 1 week after childbirth [19,20].

---

**Figure 1.** Approach for identifying system redesign opportunities.
During prenatal care, almost all pregnant women carry the *Boshi Kenko Techo*, a paper-based maternal and child health (MCH) handbook. The MCH handbook is filled and reviewed by the prenatal care providers. The MCH handbook contains information about the woman’s pregnancy and the child’s development and health [20,21].

Japanese women can receive midwife-led (MW-led) prenatal care or obstetrician-led (OB-led) prenatal care. Previous studies have found that pregnant women in the MW-led care group gave higher ratings to their care satisfaction and their perception of woman-centered care [22]. These results highlight the different responsibilities of the obstetricians and the midwives. The obstetricians’ focus is mostly biomedical, whereas the midwives’ focus is to promote self-care.

In Japan, the adoption of EMR systems has been steadily increasing since 2005 [23]. By 2020, the EMR adoption rate is expected to reach 90% for general hospitals [24]. Although EMR systems are regularly used in Japanese prenatal care, little is known about their use and about the attitudes of the prenatal care providers regarding them.

**Redesigning from Appropriation**

In a proposed model for artifact study, Fleming talked about the function of an artifact being one of its 5 basic properties [25]. He proposed that the function “embraces both the uses (intended functions) and the roles (unintended functions) of the object in its culture.” He also noted that functional analysis would have to involve the discussion of the human and their artifact-associated behavior.

In a similar vein, multiple studies in computer-supported cooperative work shed light on these unintended functions through the concept of appropriation. Once deployed in their contexts of use, artifacts or technologies are appropriated by their users [26,27]. Appropriation is “the way technologies are adopted, adapted and incorporated into working practice” [28]. Dourish presented the concept of appropriation as a broader view of customization, one that includes users “making use of the technologies for purposes beyond those for which it was originally designed, or to serve new ends” [28]. In this sense, Dourish noted that appropriation lies at the intersection of workplace studies and design and that understanding how technologies are appropriated is critical to developing them [28].

It has been suggested that understanding the ways in which a technology is appropriated is important to improve its design process [29,30]. Carroll argues that the appropriation of technologies is part of their design process. As the users appropriate a technology, they play a crucial role in completing its design [17]. Carroll also proposed improving the technologies’ design by harvesting the users’ needs from their appropriation activities. By deriving requirements from the appropriated technology, the designer would “design from appropriation” and involve the users as co-designers in an evolutionary design approach [17]. On a similar note, Fischer described the “impossibility of complete coverage” as one of the biggest design challenges for designing high-functionality environments. To address this challenge, he proposed “viewing the systems as open-ended and continuously adapted by the people who use them in their day-to-day work” [31].

In their work on persuasive technologies, Krischkowsky et al argued that the study of the unintended uses of a technology is critical to counteract undesired consequences [32]. Furthermore, we can find various works in the fields of human-computer interaction [31] and design [33] that aim to understand the appropriation of technologies in their contexts of use.

**Roles of Electronic Medical Record Systems**

Chase et al [34] examined the roles of electronic health record (EHR) systems regarding the collaboration between care providers. They identified 4 general EHR roles:

- Repository: the EHR allows the providers to have all the needed data in one place.
- Messenger: the EHR enables information transfer between the providers.
- Orchestrator: the EHR ensures that the right person is doing the right thing at the right time.
- Monitor: the EHR allows the identification of gaps in treatment and provides a benchmark for performance evaluation.

Through these roles, they described the ways in which the EHR system supports or hinders collaboration. Unlike this study, the roles were not extracted with the aim of redesigning from appropriation and, thus, were not translated into design recommendations.

**Methods**

**Overview**

In this study, we used mixed research methods in an exploratory sequential manner—a qualitative study followed by a quantitative study. We used the results from the qualitative method to inform the quantitative method. The sequence of the applied methods is shown in Figure 2.

After familiarizing ourselves with the prenatal care process, we conducted a field-based observational study at a Japanese prenatal care clinic. We analyzed the data following a thematic analysis approach to identify the different situated roles that the EMR system plays. Then, we administered a survey to the prenatal care staff to understand their experiences and attitudes regarding the situated roles of the EMR system. Finally, we analyzed the survey data and categorized the situated roles of the EMR system based on the users’ experiences and attitudes. In the following sections, we describe in detail how we conducted each step of the study.
Literature Review—Familiarization With the Prenatal Care Process

To rapidly gather a large corpus of knowledge and gain an initial understanding of the prenatal care process, we conducted a review targeting the existing literature on the prenatal care process and guidelines for obstetrical practices in Japan [18,19,20,21,22]. We validated our initial understanding of the process by discussing it with a practicing obstetrician at the prenatal care clinic.

Identifying the Situated Roles of the Electronic Medical Record System

Data Collection: Observational Field Study

One researcher observed a team of obstetricians and midwives providing prenatal care services at an outpatient clinic in a Japanese university hospital. In the observed clinic, a total of 5 obstetricians and 10 midwives provide prenatal care services. After obtaining the approval of 3 obstetricians to observe checkups during their shifts, the researcher conducted the observations by visiting the prenatal care outpatient clinic twice a week over a period of 3 weeks.

At the beginning of the checkups, the obstetricians explained to the pregnant women and their companions the reasons for the researcher’s presence in the clinic. The obstetricians also asked the pregnant women and their companions if they accept having the researcher observe and take notes during the checkup. After the pregnant women and their companions granted their approval, the researcher directly observed the prenatal care checkups and took notes using pen and paper. The notes contained descriptions of the interactions that took place around the EMR system, key information relating to the women’s course of pregnancy (pregnancy week, pregnancy type, and pregnancy number), sketches of the room layout and the EMR screen, and quotes and impressions from the conversations that took place around the EMR system.

In the observed clinic, there were 2 desks with computer terminals, as shown in Figure 3. One desk was used by the obstetrician and the other by the midwife. The room layout was semi-inclusive patient controlled. The pregnant women could see the EMR screen by moving the direction of their gaze [35].

During the observations, the researcher did not engage in conversations with the present parties. After the pregnant women and their companions left the clinic, the researcher asked the clinical staff questions to clarify certain occurrences. The researcher asked for information about the software that the staff used in addition to the EMR software. The researcher also asked for explanations as to why certain things were done in certain ways, for example, (1) using an image snipping tool, (2) bolding and changing the color of certain text, or (3) copying information from EMR notes and pasting them in other notes. In addition, the researcher inquired about artifacts that were used during the checkups, such as reference books that the staff used and paper files that the pregnant women exchanged with the staff.

In total, the researcher observed a team of 3 obstetricians and 6 midwives performing 37 prenatal care checkups for 35 different pregnant women between the eighth and 33rd week of their pregnancy.

Data Analysis: Thematic Analysis

After each observation, the field notes were transcribed and imported into QDA Miner, a qualitative data analysis tool made by Provalis Research [36]. After the observations were completed, the data were analyzed by 3 researchers following the 6 phases of thematic analysis described by Braun et al [37]:

- Familiarization with the data: In the beginning of the analysis process, we read and discussed the data multiple times to familiarize ourselves with it.
- Coding the data: The coding process was conducted over 3 iterations in which the codes were extended and refined. The process is described below.

While familiarizing ourselves with the data, we found that the interactions with the EMR system fall into 4 main categories: (1) interactions that support the communication, (2) interactions that hinder the communication, (3) interactions that support the clinical process, and (4) interactions that hinder the clinical process. These categories were mapped into 4 initial codes and were used in the first coding iteration.

Figure 2. The applied methods.
After the first coding iteration, we noted more specifically how the EMR system supports/hinders the communication/process. These more detailed descriptions were used to code the data in the second coding iteration.

After the second coding iteration, we extended the codes to reflect aspects that could not be captured in the original codes and we merged codes together when their contents overlapped. Using these extended and refined codes, we conducted our third coding iteration.

• Searching for the themes: after the coding was completed, 2 researchers examined the codes to see which ones could fit together under one theme. A theme was considered to be any set of codes that captures a significant or interesting unintended way in which the parties interact with the EMR system.

• Reviewing the themes: in this step, we discussed which themes qualify as situated roles of the EMR system. A situated role of the EMR system would be any theme that reflects an unintended way that the users engage with, relate to, and perceive the system in its context of use. When deciding which themes to keep and which themes to discard, we answered the following questions:
  1. Does the theme really reflect an unintended way that the users engage with, relate to, and perceive the system?
  2. Does the theme make sense?
  3. Does the data that we collected support our conclusion?

• Defining and naming the themes: after reviewing the themes, we finally named and clearly defined them to reflect situated roles of the EMR system.

• Producing the report: the situated roles of the EMR system are presented in the Results section.

Identifying the Attitudes of the Prenatal Care Providers Regarding the Situated Roles of the Electronic Medical Record System

Survey Design

After the situated roles were defined and named through the thematic analysis process, we administered a survey for all the prenatal care staff working at the observed clinic. In total, 15 surveys were sent out to 5 obstetricians and 10 midwives. The survey participants included the obstetricians and midwives that we observed in the field study.

First, the purpose of the survey was to validate the situated roles through the experiences of the users. Second, we wanted to understand how frequently the users want to experience the EMR roles and how important is each role to them. Therefore, for each situated role, we asked 3 questions:

1. Currently, how frequently do you experience [role’s definition]?
2. Optimally, how frequently would you experience [role’s definition]?
3. It is important to me that the EMR system does [role’s definition] OR It is important to me that the EMR system does not [role’s definition]

The purpose of the first question was to validate the situated roles through the users’ current experiences. Answers to this type of question were reported using a 6-point Likert scale ranging from very frequently (1) to not at all (6).

The purpose of the second question was to understand how frequently the users wanted to experience each role. Answers to this type of question were reported using a 6-point Likert scale ranging from very frequently (1) to not at all (6).

The purpose of the third question was to understand the importance of each situated role. These statements were formulated based on the roles’ nature. For favorable roles, we asked about the importance of their presence. For unfavorable roles, we asked about the importance of their absence. Answers to these statements were reported using a 4-point Likert scale ranging from strongly agree (1) to disagree (4).

In addition to the questions regarding the situated roles, the respondents were asked for their job title and the number of years they had used the target EMR system.

The survey was designed over several iterations. It was pretested by researchers in medical informatics and human-computer interaction who evaluated the structure, understandability, scales,
and formulation of the questions. Finally, the survey was pilot tested with 2 graduate students and refined based on their experience and feedback. The final survey was administered in Japanese.

**Survey Participants**

We received a total of 15 survey responses, 5 from obstetricians and 10 from midwives. Table 1 shows the experience of the participants with the EMR system.

**Survey Analysis**

The purpose of the survey was threefold: (1) validating the situated roles through the experiences of the users, (2) understanding how often the users want to experience each situated role, and (3) understanding how important each situated role is to the users.

**Validation of the Situated Roles**

On the basis of the survey responses, we considered that a situated role is validated if at least one respondent reports experiencing it occasionally.

We also categorized the situated roles into 3 categories reflecting the extent to which they are currently experienced by the users:

1. **Frequently:** more than half of the respondents experienced the role at least frequently.
2. **Occasionally:** more than half of the respondents experienced the role at least occasionally.
3. **Rarely:** more than half of the respondents experienced the role rarely at most.

**Desired Frequency of the Situated Roles**

We categorized the situated roles into 3 categories reflecting the extent to which they are wanted to be experienced by the users:

1. **Frequently:** more than half of the respondents wanted the role to be experienced at least frequently.
2. **Occasionally:** more than half of the respondents wanted the role to be experienced at least occasionally.
3. **Rarely:** more than half of the respondents wanted to experience the role rarely at most.

**Importance of the Situated Roles**

On the basis of the survey responses, we categorized the situated roles into 4 categories reflecting their degree of importance for the users:

1. **Very important:** at least half of the respondents strongly agree that the presence or absence of the role is important.
2. **Important:** at least half of the respondents agree that the presence or absence of the role is important.
3. **Somewhat important:** at least half of the respondents somewhat agree that the presence or absence of the role is important.
4. **Not important at all:** at least half of the respondents disagree that the presence or absence of the role is important.

Table 1. The survey participants.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Job</th>
<th>Experience with the EMR system (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>Obstetrician</td>
<td>1</td>
</tr>
<tr>
<td>O2</td>
<td>Obstetrician</td>
<td>1</td>
</tr>
<tr>
<td>O3</td>
<td>Obstetrician</td>
<td>7</td>
</tr>
<tr>
<td>O4</td>
<td>Obstetrician</td>
<td>4</td>
</tr>
<tr>
<td>O5</td>
<td>Obstetrician</td>
<td>7</td>
</tr>
<tr>
<td>M1</td>
<td>Midwife</td>
<td>12</td>
</tr>
<tr>
<td>M2</td>
<td>Midwife</td>
<td>3</td>
</tr>
<tr>
<td>M3</td>
<td>Midwife</td>
<td>12</td>
</tr>
<tr>
<td>M4</td>
<td>Midwife</td>
<td>3</td>
</tr>
<tr>
<td>M5</td>
<td>Midwife</td>
<td>5</td>
</tr>
<tr>
<td>M6</td>
<td>Midwife</td>
<td>13</td>
</tr>
<tr>
<td>M7</td>
<td>Midwife</td>
<td>1</td>
</tr>
<tr>
<td>M8</td>
<td>Midwife</td>
<td>5</td>
</tr>
<tr>
<td>M9</td>
<td>Midwife</td>
<td>3</td>
</tr>
<tr>
<td>M10</td>
<td>Midwife</td>
<td>6</td>
</tr>
</tbody>
</table>

*EMR: electronic medical record.*
Results

Situated Roles of the Electronic Medical Record System

In total, we were able to extract 10 distinct situated roles that the EMR system plays in prenatal care settings.

We found 4 situated roles relating to the communication between the providers, the pregnant women, and their companions, namely: (1) the wingman, (2) the accomplice, (3) the third wheel, and (4) the bouncer.

Regarding the clinical process, we found that the EMR system plays 6 different situated roles, namely: (1) the messenger, (2) the summarizer, (3) the assistant, (4) the gossip, (5) the alien, and (6) the bureaucrat. Table 2 shows the situated roles and their definitions.

The Wingman

As a wingman, the EMR system supports the care providers in the explanation process.

During the checkups, the clinical staff verbally communicated clinical information to the pregnant women and their companions. This communication helps the pregnant women and their companions understand the current state of the pregnancy and the logic behind clinical decisions. In the observations, the obstetricians used the EMR system as a support tool to provide clinical information and explanations. We observed the obstetricians pointing toward the screen while reading their EMR notes and explaining them. The obstetricians also used automatically generated charts and ultrasound images from their EMR notes to visually communicate information to the pregnant women and their companions.

However, the obstetricians did not always automatically employ this strategy. In one case, while the obstetrician was explaining, the pregnant woman started leaning toward the EMR system’s screen to see the image that the obstetrician was looking at. Only after realizing that the pregnant woman was interested in seeing the image did the obstetrician rotate the monitor in her direction.

The Accomplice

As an accomplice, the EMR system helps pause communication with the pregnant women.

We found that the obstetricians used the EMR system as a tool to pause communication with the pregnant women, a strategy which proved particularly useful when their workload was high or in highly emotional situations.

One of the obstetricians expressed the need for a moment to think in which they do not have to maintain a conversation with the pregnant women. In such cases, the EMR system served as a tool to pause the conversation and provide them with the needed moment to think.

Moreover, talking about pregnancies, especially complicated ones, could result in highly emotional situations. In this case, the EMR system provided the obstetricians with a bubble allowing them to distance themselves from the interaction. In one observed case, the obstetrician had to tell the pregnant woman that her pregnancy must be terminated. This woman had already experienced a pregnancy termination. After receiving the information, the pregnant woman started crying. At that moment, the obstetrician resorted to the EMR system to avoid looking at the pregnant woman and allow her to privately wipe her tears and stop herself from crying. When the obstetrician turned to the EMR system, the midwife left her desk and went toward the pregnant woman with a tissue box in hand. The midwife continued standing next to the pregnant woman while the obstetrician was working on the EMR system.

The Third Wheel

As a third wheel, the EMR system distracts the care providers from communicating with the pregnant women.

We found that the obstetricians spent a major part of the checkup time keyboarding and facing the EMR screen. During the obstetricians’ data input time, the pregnant women waited silently in their chair, looked closely at the EMR screen to see what their obstetrician was typing, or tried to initiate a conversation with the obstetrician or their companions.

The Bouncer

As a bouncer, the EMR system excludes the pregnant women and their companions from the electronic medical record.

We found that the obstetricians used the EMR system as a barrier to exclude the pregnant women and their companions from the electronic medical record.
While inputting data, the obstetricians responded to the pregnant women in various ways. Most of the time, they responded by turning their heads slightly away from the screen toward the pregnant woman. When the pregnant woman continued to ask questions or tried to engage in conversation, the obstetricians either started to alternate quickly between the screen and her or stopped inputting data and turned their chair away from the desk to face and respond to her. In some cases, they fully rotated their chair, but in most cases, they turned it halfway between their desk and the pregnant woman.

**The Bouncer**

As a bouncer, the EMR system creates an exclusive environment by physically excluding the pregnant women and their companions.

On multiple occasions, we found that the pregnant women and their companions showed interest in looking at the EMR. However, the pregnant women had to actively get closer to the screen while their companions’ assigned chairs were placed too far from the screen, leading most of them to stop trying to look at the screen after a while.

On one occasion, the companion of the pregnant woman stood up to get a better view of the EMR screen. After standing up and realizing that he still cannot get a clear view, he tilted his head and body forward in the direction of the screen. When he realized that, even in this position, he cannot clearly see the contents of the EMR, he went back to his seat. After some time, he got up again, tilted forward toward the screen and went back to his seat, clearly feeling disappointed. A while later, he repeated the same sequence: he stood up, tilted forward, and sat down again. After sitting down, he gazed at the floor, bored and frustrated. Finally, he stood up, moved closer to the pregnant woman and to the EMR screen and remained standing there until the end of the checkup.

**The Messenger**

As a messenger, the EMR system enables the communication of information between the care providers.

In the case of the observed clinic, every pregnancy was cared for by multiple obstetricians and midwives. The rotation of the clinical staff required them to communicate the pregnant women’s health data. The EMR system was the main tool for communicating clinical information to ensure continuity of care. The EMR system, in this case, provided seamless communication between the clinical staff over time and staff rotations.

Conversely, when pregnant women were transferred from other clinics, the team only had access to the paper records that they had brought with them. In this case, the team created new EMRs for the women. However, the previous notes existing in the paper records were not transferred to the newly created EMRs.

Even though the pregnant women and their family members are involved in communicating health-related information to the care providers, they did not have the ability to directly add information into the EMR. During the checkups, through conversations with the pregnant women and notes from the women’s MCH handbooks, the providers gathered information and added them into EMR memos. However, what went into the EMR remained under the full control of the care providers.

In one examination, a pregnant woman, with a history of high blood pressure, brought along a paper containing a list of blood pressure measurements that she self-monitored and recorded. The obstetrician reviewed the measurements and handed the paper back to the woman. Then, the obstetrician wrote a note in an EMR memo regarding the measurements. However, the full list of blood pressure measurements remained out of the woman’s EMR.

**The Summarizer**

As a summarizer, the EMR system provides the care providers with a summary of the pregnancy’s current state and care course.

The EMR system allows the prenatal care providers to have all the health information in one place. On multiple occasions, before calling a pregnant woman into the clinic, the obstetricians quickly navigated through the previous EMR notes to form a mental summary of her current course of pregnancy.

However, the EMR system did not allow for a quick understanding of the current state of the pregnancy. One obstetrician noted, “we would like to see the course of care in one glance. With paper records, it was easier to do that. However, with this system, it takes a lot of clicking and scrolling to get the full image.”

The staff needed the EMR system to act as a summarizer. To achieve that, the obstetricians employed a workaround. To give themselves and the other providers a quick understanding of the care course, the obstetricians emphasized certain parts of their EMR notes by changing the size, boldness, and color of the text.

**The Assistant**

As an assistant, the EMR system facilitates the management and preparation of the checkups.

At the beginning of their shift, using the EMR system, the obstetrician and the midwife viewed the list of scheduled checkups. Knowing the number of checkups, they could estimate the workload for the day. Based on that information, they adapted the speed of their work and the duration of checkups. Moreover, using the scheduled checkups list, the obstetrician and midwife knew who they were examining next and had access to her records before the checkup. Before calling the woman in, they reviewed the previous notes and discussed the current state of the woman’s pregnancy. Using this information, they could form a picture of what care actions they needed to perform once the pregnant woman was called in. By allowing for previous preparation, the EMR system makes the checkups run more smoothly. It eliminates the need for the staff to orient themselves and for the pregnant woman to explain the reason for her visit at the beginning of her visit.

**The Gossip**

As a gossip, the EMR system is not completely trusted with sensitive information. In our analysis, we found that the clinical staff hesitate to include highly sensitive information in the pregnant women’s records because of privacy and legal
concerns. In one of our discussions with the staff, one midwife stated,

> If we have concerns over some psychosocial issues such as domestic abuse, we note it indirectly in the record. We do not write it literally; we use codes to pass the message to the other clinical staff.

Employing this sort of strategy to document sensitive information implies that the EMR system is not completely trusted by the staff with information that is usually considered private or could be used for legal purposes.

**The Alien**

As an alien, the EMR system (1) has low learnability, (2) requires a high level of recall, and (3) has an interface that is not optimized for routine tasks.

The difficulty of learning how to use the EMR system was one of the problems noted by the obstetricians. One obstetrician mentioned that it took them at least 1 month to get used to the system. Moreover, during a checkup, a staff member walked into the clinic and asked the obstetrician a question regarding the use of the EMR system, which the obstetrician answered by guiding them through the interface.

Furthermore, the EMR system appeared to require a high level of memory recall. The obstetricians frequently paused and tried to recall in which tab a specific field, or note, was placed. The inconvenience of the manual data input was further amplified by an interface design that was not optimized for routine data input and data retrieval tasks.

In addition, the EMR system suffered from performance-related issues. In certain cases, the system would temporarily stop responding or have a slow response time. These issues occurred particularly when the providers opened a new EMR. Even though 90% of the common database queries are usually cached, and the list of patients is previously compiled, opening a new EMR required more than 10 seconds in certain observed cases. This poor performance resulted in obvious frustration and time loss. To counter this issue, some midwives employed a sort of manual caching where they anticipated the need to open the records, opened the records, and let them load before they actually needed them.

**The Bureaucrat**

As a bureaucrat, the EMR system requires the care providers to halt the care process to input data.

During the checkups, the providers continuously collect data from multiple sources and add it into the EMR. Those sources include conversations with the pregnant women and their family members; ultrasound imaging devices; and measuring devices such as blood pressure meters, weighing scales, and measuring tapes. The lack of integration between the medical devices and the EMR system required the providers to manually input most of the data that they collect. To do so, they had to intermittently pause their clinical flow. Below are some examples of occurrences encountered during the observations.

Before entering the clinic, the pregnant women use a blood pressure meter and a weighing scale located in the waiting room.

The machines print the measurements on small paper receipts. Once they enter the clinic, the women hand the paper receipts and their MCH handbooks to the midwives. During the checkup, the midwives copy the measurements into the MCH handbook and then input them into the EMR. The process of copying the data could take up to 3 min. After they copy the measurements, the midwives throw the small paper receipts in a trash bin under their desks. On 2 different occasions, during ongoing checkups, the midwives had to look in the trash bin for receipts that they had previously thrown away. In one of those occasions, a nurse had to come in, put gloves on, and help the midwife look inside the trash bin.

In addition, the midwives routinely measure the belly circumference before the obstetricians start to conduct the ultrasound. Using a measurement tape, they measure the belly twice, vertically and then horizontally. After the second measurement, the midwives sometimes retake the first measurement, as that they might have forgotten the first measure. Once they finish measuring, some midwives prepare the women for the ultrasound, turn off the lights and then head back to their desks to input the measures. As this increases the risk of forgetting the measures, other midwives prefer to head fast to their desks, input the measures in the MCH handbook and the EMR, and then return to the woman to prepare her for the ultrasound.

As for the obstetricians, they routinely use 2 ultrasound devices to collect data. After they finish conducting the ultrasounds, they reflect on the results and summarize them inside free-text EMR notes. Then, they add the ultrasound images to the notes. To do so, they manually copy the information from the output of the ultrasound devices into the EMR. In addition, they use an image snipping tool to take screenshots of the ultrasound images and then they paste the images inside the EMR notes.

It is important to note that similarly to the third wheel, this role is manifested when the providers input data into the EMR system. However, as a third wheel, the EMR system hinders the communication between the providers and the pregnant women, whereas as a bureaucrat, the EMR system hinders the clinical workflow.

**Attitudes of the Prenatal Care Providers Regarding the Situated Roles of the Electronic Medical Record System**

All the prenatal care staff responded to the survey. In total, we received 15 responses, 10 responses from midwives and 5 responses from obstetricians.

**Validation of the Situated Roles**

To validate the situated roles, we analyzed the responses regarding the users’ current experience. We looked at the midwives’ and obstetricians’ answers separately. All the situated roles were validated as at least one respondent reported experiencing them occasionally. Tables 3 and 4 show the responses of the obstetricians and midwives, respectively. The numbers in the table indicate the number of respondents that chose the option.
To better understand the extent to which different roles are experienced, we assigned them to 3 different categories:

1. Frequently: more than half of the respondents experienced the role at least frequently.
2. Occasionally: more than half of the respondents experienced the role at least occasionally.
3. Rarely: more than half of the respondents experienced the role rarely at most.

The extent to which the respondents experience the situated roles is shown in Table 5.

Table 3. The current frequency of experiencing the situated roles as reported by the obstetricians. Participant identifier listed in parentheses.

<table>
<thead>
<tr>
<th>Situated role</th>
<th>Very frequently (n)</th>
<th>Frequently (n)</th>
<th>Occasionally (n)</th>
<th>Rarely (n)</th>
<th>Very rarely (n)</th>
<th>Not at all (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wingman</td>
<td>0</td>
<td>2 (O3, O5)</td>
<td>2 (O1, O4)</td>
<td>1 (O2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Accomplice</td>
<td>0</td>
<td>2 (O1, O5)</td>
<td>0</td>
<td>0</td>
<td>2 (O3, O4)</td>
<td>1 (O2)</td>
</tr>
<tr>
<td>Third wheel</td>
<td>0</td>
<td>0</td>
<td>3 (O1, O2, O4)</td>
<td>0</td>
<td>0</td>
<td>2 (O3, O5)</td>
</tr>
<tr>
<td>Bouncer</td>
<td>0</td>
<td>0</td>
<td>1 (O4)</td>
<td>0</td>
<td>2 (O1, O3)</td>
<td>2 (O2, O5)</td>
</tr>
<tr>
<td>Messenger</td>
<td>1 (O3)</td>
<td>3 (O2, O4, O5)</td>
<td>1 (O1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Summarizer</td>
<td>2 (O2, O3)</td>
<td>3 (O1, O4, O5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Assistant</td>
<td>0</td>
<td>3 (O1, O2, O3)</td>
<td>1 (O5)</td>
<td>0</td>
<td>0</td>
<td>1 (O4)</td>
</tr>
<tr>
<td>Gossip</td>
<td>0</td>
<td>0</td>
<td>3 (O1, O4, O5)</td>
<td>1 (O2)</td>
<td>1 (O3)</td>
<td>0</td>
</tr>
<tr>
<td>Alien</td>
<td>0</td>
<td>0</td>
<td>4 (O1, O3, O4, O5)</td>
<td>1 (O2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bureaucrat</td>
<td>0</td>
<td>2 (O1, O5)</td>
<td>2 (O2, O3)</td>
<td>0</td>
<td>1 (O4)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. The current frequency of experiencing the situated roles as reported by the midwives. Participant identifier listed in parentheses.

<table>
<thead>
<tr>
<th>Situated role</th>
<th>Very frequently (n)</th>
<th>Frequently (n)</th>
<th>Occasionally (n)</th>
<th>Rarely (n)</th>
<th>Very rarely (n)</th>
<th>Not at all (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wingman</td>
<td>0</td>
<td>1 (M8)</td>
<td>3 (M1, M6, M7)</td>
<td>2 (M4, M5)</td>
<td>1 (M2)</td>
<td>3 (M3, M9, M10)</td>
</tr>
<tr>
<td>Accomplice</td>
<td>0</td>
<td>0</td>
<td>7 (M1, M3, M4, M5, M6, M7, M8)</td>
<td>1 (M2)</td>
<td>2 (M9, M10)</td>
<td>0</td>
</tr>
<tr>
<td>Third wheel</td>
<td>1 (M3)</td>
<td>0</td>
<td>3 (M1, M4, M6)</td>
<td>2 (M2, M5)</td>
<td>1 (M8)</td>
<td>3 (M7, M9, M10)</td>
</tr>
<tr>
<td>Bouncer</td>
<td>0</td>
<td>1 (M3)</td>
<td>6 (M1, M4, M5, M8, M9, M10)</td>
<td>3 (M2, M6, M7)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Messenger</td>
<td>4 (M2, M3, M4, M7)</td>
<td>6 (M1, M5, M6, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Summarizer</td>
<td>4 (M3, M4, M6, M7)</td>
<td>6 (M1, M2, M5, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Assistant</td>
<td>5 (M1, M3, M4, M7, M9)</td>
<td>4 (M2, M6, M8, M10)</td>
<td>1 (M5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gossip</td>
<td>0</td>
<td>0</td>
<td>1 (M8)</td>
<td>1 (M10)</td>
<td>7 (M1, M2, M4, M5, M6, M7, M9)</td>
<td>1 (M3)</td>
</tr>
<tr>
<td>Alien</td>
<td>0</td>
<td>5 (M3, M4, M5, M9, M10)</td>
<td>4 (M1, M2, M7, M8)</td>
<td>1 (M6)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bureaucrat</td>
<td>1 (M3)</td>
<td>1 (M1)</td>
<td>6 (M2, M4, M5, M6, M8, M9)</td>
<td>1 (M7)</td>
<td>1 (M10)</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5. The extent to which the respondents experience the situated roles.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Obstetricians</th>
<th>Midwives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequently</td>
<td>Summarizer; messenger; assistant</td>
<td>Summarizer; messenger; assistant; alien</td>
</tr>
<tr>
<td>Occasionally</td>
<td>Alien; wingman; third wheel; gossip</td>
<td>Accomplice; bureaucrat; bouncer</td>
</tr>
<tr>
<td>Rarely</td>
<td>Accomplice; bouncer; bureaucrat</td>
<td>Wingman; third wheel; gossip</td>
</tr>
</tbody>
</table>
Desired Frequency of the Situated Roles

To understand the desired frequency of each situated role, we analyzed the users’ responses for how frequently they would like to experience the roles. We looked at the midwives’ and obstetricians’ answers separately.

Tables 6 and 7 show the responses of the obstetricians and midwives respectively. The numbers in the table indicate the number of respondents that chose the option.

To better understand the extent to which different situated roles are wanted, we assigned them to 3 different categories:
1. Frequently: more than half of the respondents wanted the role to be experienced at least frequently.
2. Occasionally: more than half of the respondents wanted the role to be experienced at least occasionally.
3. Rarely: more than half of the respondents wanted to experience the role rarely at most.

The extent to which the respondents want to experience the situated roles is shown in Table 8.

Table 6. The desired frequency of experiencing the situated roles as reported by the obstetricians. Participant identifier listed in parentheses.

<table>
<thead>
<tr>
<th>Situated role</th>
<th>Very frequently (n)</th>
<th>Frequently (n)</th>
<th>Occasionally (n)</th>
<th>Rarely (n)</th>
<th>Very rarely (n)</th>
<th>Not at all (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wingman</td>
<td>0</td>
<td>3 (O2, O3, O5)</td>
<td>2 (O1, O4)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Accomplice</td>
<td>0</td>
<td>0</td>
<td>2 (O1, O5)</td>
<td>0</td>
<td>0</td>
<td>3 (O2, O3, O4)</td>
</tr>
<tr>
<td>Third wheel</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 (O2)</td>
<td>0</td>
<td>4 (O1, O3, O4, O5)</td>
</tr>
<tr>
<td>Bouncer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5 (O1, O2, O3, O4, O5)</td>
</tr>
<tr>
<td>Messenger</td>
<td>2 (O2, O3)</td>
<td>3 (O1, O4, O5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Summarizer</td>
<td>2 (O2, O3)</td>
<td>3 (O1, O4, O5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Assistant</td>
<td>0</td>
<td>5 (O1, O2, O3, O4, O5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gossip</td>
<td>0</td>
<td>0</td>
<td>3 (O1, O4, O5)</td>
<td>1 (O2)</td>
<td>1 (O3)</td>
<td>0</td>
</tr>
<tr>
<td>Alien</td>
<td>0</td>
<td>0</td>
<td>2 (O4, O5)</td>
<td>2 (O2, O3)</td>
<td>1 (O1)</td>
<td>0</td>
</tr>
<tr>
<td>Bureaucrat</td>
<td>0</td>
<td>0</td>
<td>1 (O5)</td>
<td>1 (O2)</td>
<td>1 (O1)</td>
<td>2 (O3, O4)</td>
</tr>
</tbody>
</table>

Table 7. The desired frequency of experiencing the situated roles as reported by the midwives. Participant identifier listed in parentheses.

<table>
<thead>
<tr>
<th>Situated role</th>
<th>Very frequently (n)</th>
<th>Frequently (n)</th>
<th>Occasionally (n)</th>
<th>Rarely (n)</th>
<th>Very rarely (n)</th>
<th>Not at all (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wingman</td>
<td>0</td>
<td>2 (M2, M5)</td>
<td>5 (M1, M3, M4, M6, M7)</td>
<td>1 (M8)</td>
<td>0</td>
<td>2 (M9, M10)</td>
</tr>
<tr>
<td>Accomplice</td>
<td>0</td>
<td>1 (M8)</td>
<td>4 (M1, M3, M6, M7)</td>
<td>2 (M4, M5)</td>
<td>2 (M2, M9)</td>
<td>1 (M10)</td>
</tr>
<tr>
<td>Third wheel</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 (M5)</td>
<td>9 (M1, M2, M3, M4, M6, M7, M8, M9, M10)</td>
</tr>
<tr>
<td>Bouncer</td>
<td>0</td>
<td>1 (M1)</td>
<td>1 (M8)</td>
<td>1 (M7)</td>
<td>1 (M2)</td>
<td>6 (M3, M4, M5, M6, M9, M10)</td>
</tr>
<tr>
<td>Messenger</td>
<td>5 (M2, M3, M4, M5, M7)</td>
<td>5 (M1, M6, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Summarizer</td>
<td>8 (M2, M3, M4, M5, M6, M7, M8, M10)</td>
<td>2 (M1, M9)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Assistant</td>
<td>5 (M1, M2, M3, M4, M7)</td>
<td>4 (M6, M8, M9, M10)</td>
<td>1 (M5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gossip</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6 (M1, M6, M7, M8, M9, M10)</td>
<td>4 (M2, M3, M4, M5)</td>
</tr>
<tr>
<td>Alien</td>
<td>0</td>
<td>0</td>
<td>2 (M4, M7)</td>
<td>4 (M1, M2, M6, M8)</td>
<td>4 (M3, M5, M9, M10)</td>
<td>0</td>
</tr>
<tr>
<td>Bureaucrat</td>
<td>1 (M8)</td>
<td>0</td>
<td>1 (M6)</td>
<td>3 (M1, M4, M7)</td>
<td>2 (M2, M9)</td>
<td>3 (M3, M5, M10)</td>
</tr>
</tbody>
</table>
We assumed that favorable roles are ones that users want to experience more frequently than they currently do. Unfavorable roles are ones that users want to experience less frequently than they currently do. Accordingly, the summarizer, the messenger, the assistant, and the wingman were regarded as favorable by both obstetricians and midwives. However, the obstetricians wanted to experience the wingman more frequently than the midwives. The alien, the third wheel, the bureaucrat, and the bouncer were regarded as unfavorable by both groups.

As for the gossip and accomplice, we considered them to be ambivalent. Their ideal frequency is the same as their current frequency. The midwives considered the gossip unfavorable, but the obstetricians wanted to experience it occasionally. Conversely, the obstetricians considered the accomplice unfavorable, but the midwives wanted to experience it occasionally.

**Importance of the Situated Roles**

As mentioned in the Methods section, for favorable roles, we asked the staff about the importance of their presence. For unfavorable roles, we asked about the importance of their absence. The accomplice role was treated as favorable to the prenatal care staff as it is created by them. The gossip role was treated as unfavorable as it denotes a lack of trust in the system.

To understand the importance of the situated roles, we looked at how much the respondents agreed that the roles’ presence or absence was important. We also looked at the midwives’ and obstetricians’ answers separately. Tables 9 and 10 show the responses of the obstetricians and midwives, respectively. The numbers in the table indicate the number of respondents that chose the option.

To better understand the importance of the different situated roles, we assigned them to 4 different categories:

1. **Very important**: at least half of the respondents strongly agree that the presence or absence of the role is important.
2. **Important**: at least half of the respondents agree that the presence or absence of the role is important.
3. **Somewhat important**: at least half of the respondents somewhat agree that the presence or absence of the role is important.
4. **Not important at all**: at least half of the respondents disagree that the presence or absence of the role is important.

The situated roles’ levels of importance are shown in Table 11.
Table 10. The extent to which the midwives agree that the situated roles are important. Participant identifier listed in parentheses.

<table>
<thead>
<tr>
<th>Situated role</th>
<th>Strongly agree (n)</th>
<th>Agree (n)</th>
<th>Somewhat agree (n)</th>
<th>Disagree (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of wingman</td>
<td>1 (M5)</td>
<td>7 (M1, M2, M3, M4, M6, M7, M8)</td>
<td>2 (M9, M10)</td>
<td>0</td>
</tr>
<tr>
<td>Absence of accomplice</td>
<td>0</td>
<td>4 (M3, M5, M6, M8)</td>
<td>6 (M1, M2, M4, M7, M9, M10)</td>
<td>0</td>
</tr>
<tr>
<td>Absence of third wheel</td>
<td>8 (M1, M2, M3, M4, M5, M6, M8, M10)</td>
<td>2 (M7, M9)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Absence of bouncer</td>
<td>0</td>
<td>5 (M1, M2, M3, M5, M6)</td>
<td>3 (M4, M7, M8)</td>
<td>2 (M9, M10)</td>
</tr>
<tr>
<td>Presence of messenger</td>
<td>10 (M1, M2, M3, M4, M5, M6, M7, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Presence of summarizer</td>
<td>10 (M1, M2, M3, M4, M5, M6, M7, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Presence of assistant</td>
<td>5 (M1, M2, M3, M4, M7)</td>
<td>5 (M5, M6, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Absence of gossip</td>
<td>6 (M1, M2, M3, M4, M5, M6, M7, M8, M9, M10)</td>
<td>4 (M6, M7, M8, M10)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Absence of alien</td>
<td>10 (M1, M2, M3, M4, M5, M6, M7, M8, M9, M10)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Absence of bureaucrat</td>
<td>6 (M1, M3, M5, M6, M7, M8)</td>
<td>4 (M2, M4, M9, M10)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 11. The situated roles’ levels of importance.

<table>
<thead>
<tr>
<th>Importance</th>
<th>Obstetricians</th>
<th>Midwives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very important</td>
<td>Presence of summarizer; presence of messenger; presence of assistant; absence of alien</td>
<td>Presence of summarizer; presence of messenger; presence of assistant; absence of alien; absence of third wheel; absence of bureaucrat; absence of gossip</td>
</tr>
<tr>
<td>Important</td>
<td>Presence of wingman; absence of third wheel; absence of bureaucrat; absence of gossip</td>
<td>Presence of wingman</td>
</tr>
<tr>
<td>Somewhat important</td>
<td>Presence of accomplice</td>
<td>Presence of accomplice; absence of bouncer</td>
</tr>
<tr>
<td>Not important at all</td>
<td>Absence of bouncer</td>
<td><em>a</em></td>
</tr>
</tbody>
</table>

*aNot applicable.*

Discussion

Principal Findings

We identified 10 situated roles that EMR systems play in Japanese prenatal care. On the basis of the feedback of the obstetricians and the midwives, we validated the roles and understood the users’ wants and priorities regarding them. Our results offer multiple EMR redesign opportunities in prenatal care. Figures 4 and 5 can serve as redesign guidelines for satisfying the wants of the obstetricians and the midwives, respectively.

To improve the usability and usefulness of the EMR systems, designers can amplify the favorable roles (the roles wanted to be experienced frequently) and minimize the unfavorable roles (the roles wanted to be experienced rarely). To align their design activities with the priorities of the users, designers can focus on the roles reported as important by the users. To increase the impact of their redesigns, the designers can focus on minimizing unfavorable roles that are frequently experienced, for example, the alien, or amplifying favorable roles that are less frequently experienced, for example, the wingman.

The gossip and the accomplice roles seem to be ambivalent. The midwives want to clearly document sensitive psychosocial information in most cases, whereas the obstetricians seem to be more hesitant. Conversely, the midwives want to use the EMR system occasionally to have a moment to think, whereas obstetricians prefer to rarely do so.

Through the gossip role, our study highlighted the challenge of documenting sensitive psychosocial information in EMR systems. A reason for not documenting such information could be that psychosocial information is viewed as *subjective* and sometimes *not legitimate enough* to be placed in a permanent medical record [38]. Conversely, when such information is documented, it may be documented in vague and ambiguous ways. This behavior may be attributed to the health care providers’ distrust in the security of EMR systems or their unfamiliarity with the laws governing the access and use of health care data. It is important to note that in Japan, one in every 20 women may experience domestic violence (DV) during...
pregnancy [39]. Furthermore, pregnant Japanese women want their psychosocial information to be documented in detail in their EMRs [40]. Therefore, clearly documenting sensitive information is needed to (1) respect the preferences of the pregnant women and (2) address psychosocial issues, such as DV, that are not easily disclosed or discussed [41]. Further investigations should be conducted to understand if the gossip role is a result of the EMR system’s design as it might create a feeling of distrust for the medical staff, or if it is a result of the medical staff’s uncertainty regarding the laws governing health care data.

As for the accomplice role, it is created by the staff to pause the conversation with the pregnant women. In this case, the reasons behind this role should be further understood as it presents a possible conflict between the needs of health care providers and patients and raises an important question: whose needs should we consider and prioritize when designing EMR systems?

Our results also suggest that the EMR system is viewed first and foremost as a tool for supporting the clinical workflow. The messenger, summarizer, assistant, and alien were very important situated roles for both user groups. These results shed light on clear redesign targets, ones that improve the EMR system’s capabilities as a tool to support the clinical process.

However, the EMR system’s roles as a tool that supports the communication with the pregnant women seems secondary, as shown through the bouncer role. Both midwives and obstetricians do not think it is important that the pregnant women see the EMR screen. Conversely, they think it is important to use the EMR screen as an explanation support tool. Further research is needed to understand the attitudes of the prenatal care providers regarding the EMR system as a communication support tool.

A previous study found no difference between the perceptions of physicians and nurses regarding EMR systems [13]. Even though the experiences and wants of the midwives and the obstetricians overlapped for certain situated roles (the messenger, the summarizer, the assistant, and the alien), they differed for others. Our results imply that the users’ experiences with the EMR system, and their aspirations regarding it, are related to the nature and purpose of their job. Aligning with previous findings in Japanese prenatal care settings [22], these results highlight the different responsibilities that obstetricians and midwives have in prenatal care, as it can be argued that the obstetricians’ focus is clinical while the midwives’ focus is more psychosocial. Moreover, in OB-led checkups, the midwife assists the obstetrician instead of being the main caretaker. Consequently, the pregnant woman’s chair faces the obstetrician’s desk. The desk placement in this case might affect the situated roles exhibited by the EMR system.

Accordingly, EMR redesign efforts should take into consideration not only the prenatal care context but also the different types of EMR users in this context. One direction to explore is having 2 different EMR systems—one for the midwives and another for the obstetricians—rather than having a general prenatal care EMR system. In this case, we need to ensure that the use of different systems does not hinder collaborative care [42].

Figure 4. The situated roles of the electronic medical record system for the obstetricians.
Limitations
The first limitation of this study is the number of observations. In addition, the obstetricians that we observed were all males. As the communication style may differ substantially depending on the obstetrician and pregnant woman, more situated roles may be identified through additional observations that include female obstetricians.

Moreover, our results may not be generalizable. The readers will have to assess the degrees to which our results can be transferred to health care or cultural settings of their interest.

In other respects, it would be valuable to quantify the number of occurrences relating to a situated role. Owing to the limited number of field observations that we were able to conduct, our analysis aimed to provide a purely qualitative and nuanced description of the situated roles, without counting the frequency of occurrences. Quantifying the occurrences is particularly valuable to evaluate redesign activities. To evaluate the success of a redesign activity, designers may need to objectively quantify and compare the frequency of certain occurrences in the pre-redesign and postredesign phases. To facilitate the redesign evaluation process, our future work will look into ways to automatically and objectively quantify the occurrences without the need for manual data analysis.

Conclusions
By looking at how the prenatal care providers appropriated the EMR system, we identified 10 situated roles that EMR systems play in Japanese prenatal care settings. We also identified the wants and the priorities of the users regarding the EMR system’s situated roles.

We found that prenatal care providers mainly use the EMR system as a summarizer to have a quick summary of the pregnancy course, as a messenger to communicate patient information across staff rotations, and as an assistant to prepare for the checkups. They would like to use the EMR system as a tool to support their explanations to the pregnant women more frequently. We also found that the providers may use the EMR system as an accomplice to pause communication with the pregnant women. Interestingly, they do not think it is important for the pregnant women to view the EMR screen during the checkups. Our results further highlighted the lack of trust in the security of EMR systems and how it can negatively affect prenatal care provision.

In other respects, we found a difference in the experiences and attitudes of the obstetricians and the midwives regarding the use of the EMR system. These findings imply the need for different EMR system designs to better fit the job descriptions and tasks of the different user groups.

Our study serves as an example to show how EMR systems can be redesigned from appropriation. Our results could be used to redesign EMR systems to better fit the contextual needs of their users in prenatal care.

Acknowledgments
This research was supported by the Leading Graduate Schools Program Collaborative Graduate Program in Design, Kyoto University, Japan. The authors would like to thank Professor Kumiyo Nakakoji for her valued input and insights. The authors would also like to thank all the participants and collaborators of this project.
Conflicts of Interest
None declared.

References


Abbreviations

DV: domestic violence
EHR: electronic health record
EMR: electronic medical record
MCH: maternal and child health
MW-led: midwife-led
OB-led: obstetrician-led
Factors Associated With Electronic Health Record Usage Among Primary Care Physicians After Hours: Retrospective Cohort Study

Selasi Attipoe¹, MA; Yungui Huang², PhD, MBA; Sharon Schweikhart¹, MBA, PhD; Steve Rust², PhD; Jeffrey Hoffman²,³, MD; Simon Lin²,³, MD, MBA

¹College of Public Health, The Ohio State University, Columbus, OH, United States
²The Research Institute, Nationwide Children’s Hospital, Columbus, OH, United States
³College of Medicine, The Ohio State University, Columbus, OH, United States

*these authors contributed equally

Corresponding Author:
Selasi Attipoe, MA
College of Public Health
The Ohio State University
250 Cunz Hall
1841 Neil Avenue
Columbus, OH, 43210
United States
Phone: 1 410 908 2472
Email: selasni@hotmail.com

Abstract

Background: There is limited published data on variation in physician usage of electronic health records (EHRs), particularly after hours. Research in this area could provide insight into the effects of EHR-related workload on physicians.

Objective: This study sought to examine factors associated with after-hours EHR usage among primary care physicians.

Methods: Electronic health records usage information was collected from primary care pediatricians in a large United States hospital. Inclusion criteria consisted solely of being a primary care physician who started employment with the hospital before the study period, so all eligible primary care physicians were included without sampling. Mixed effects statistical modeling was used to investigate the effects of age, gender, workload, normal-hour usage, week to week variation, and provider-to-provider variation on the after-hour usage of EHRs.

Results: There were a total of 3498 weekly records obtained on 50 physicians, of whom 22% were male and 78% were female. Overall, more EHR usage during normal work hours was associated with decreased usage after hours. The more work relative value units generated by physicians, the more time they spent interacting with EHRs after hours ($\beta=.04, P<.001$) and overall (ie, during normal hours and after hours) ($\beta=.24, P<.001$). Gender was associated with total usage time, with females spending more time than males ($P=.03$). However, this association was not observed with after-hours EHR usage. Provider-to-provider variation was the largest and most dominant source of variation in after-hour EHR usage, which accounted for 52% of variance of total EHR usage.

Conclusion: The present study found that there is a considerable amount of variability in EHR use among primary care physicians, which suggested that many factors influence after-hours EHR usage by physicians. However, provider-to-provider variation was the largest and most dominant source of variation in after-hours EHR usage. While the results are intuitive, future studies should consider the effect of EHR use variations on workload efficiency.

(JMIR Hum Factors 2019;6(3):e13779) doi:10.2196/13779

KEYWORDS

electronic health records; health information technology; primary care physicians
**Introduction**

While some studies have suggested that electronic health records (EHRs) increase efficiency and productivity, the scientific evidence has been mixed. On workload efficiency, one study found that primary care physicians (PCPs) who used EHRs spent an extra 1.3 facetime minutes per visit and had increased their patient volume per week [1]. However, in another study, EHR implementation was found to be associated with a negative impact on productivity and efficiency in a pediatric ophthalmology clinic [2]. Moreover, systematic reviews did not demonstrate any superior productivity from the use of EHRs [3,4].

Since the introduction of EHRs, physicians have been reported to work longer hours, with many completing their data entry after clinic and at home, during evening or weekends [5]. Sinsky et al attempted to quantify “work after work” using a diary among 21 physicians [6]. Solving the scale-up and accuracy challenge of diary-based studies, Arndt et al developed a passive observation method using the access log automatically collected by EHRs so that all PCPs at an EHR site could be studied. However, little is known quantitatively about the factors associated with the usage of EHRs after-hours [7]. Identifying these factors will provide knowledge to support recommendations for interventions that could improve the user experience, enhance efficiency, and mitigate burnout. Accordingly, this retrospective study examines variations in EHR usage of PCPs, with an emphasis on after-hours EHR use.

**Methods**

**Setting**

Nationwide Children’s Hospital (NCH) is a large, free-standing US children’s hospital that has used the Epic EHR system (Epic Systems Corporation, Verona) as its enterprise-wide system since 2006. All pediatricians who generated work relative value units (wRVUs) related to ambulatory primary care activity from January 1, 2015 to June 30, 2016 were included in this study. Work RVUs are a measure of billable service volume and complexity. Clinical data, billing information, and EHR usage data were extracted from the EHRs into a database for analysis. This study was approved by the NCH’s Institutional Review Board.

**Data**

Data retrieved included physician demographics (age and gender), duration of employment (tenure), medical specialty, wRVUs generated during the study period, full-time equivalent (FTE) status, and EHR access logs. The EHR access log captures clinicians’ direct interactions with the EHR system, such as login, logout, and documentation in a patient’s chart.

We implemented the algorithm employed by Arndt et al to estimate the duration of EHR usage (ie, duration of a physician’s EHR activity) from EHR access logs [7]. This method included estimating time spent on particular activities using the EHR system’s automated event logging feature. This algorithm was validated using a time and motion study (ie, direct observation of 14 nonresident family medicine physicians by a trained medical student) [7].

EHR usage was divided into two separate time segments: normal (weekdays from 7am-6pm, using NCH workstations) and after-hours (weekdays from 6pm-7am, anytime on weekends, or anytime not using NCH workstations). All daily EHR activity durations for each physician were classified as belonging to one of these two different time segments (based on the activity timestamps) and summed for each time segment.

The main outcome variable was the duration of after-hours EHR usage. EHR access time during normal hours, along with physician age, gender, tenure, and wRVUs, were the main explanatory variables. Records for the analysis are organized into one record per physician per week, per the two time segments chosen.

**Statistical Analysis**

Mixed effect linear regression models were used to analyze data, where physician age, gender, wRVUs, and EHR usage were treated as fixed effects, and both provider-to-provider variation and week to week variation were treated as random effects. Linear mixed models were used in this analysis to account for repeated measures within the same individuals. Distributions for dependent variables were analyzed to assess the normality assumption and to determine whether transformation was needed. Analyses were all conducted using R (Version 3.3.0, 2016) and the R lme4 package.

**Results**

**Descriptive Statistics**

A total of 50 physicians, of whom 22% were male and 78% were female, met the inclusion criteria and their descriptive statistics are presented in Table 1. The study physicians generated a total of 3498 aggregated weekly records (ie, access logs). Physicians spent approximately 16 hours weekly interacting with the EHRs during their normal work hours, while spending about 3 hours weekly after-hours.

**Statistical Modeling**

Mixed effect linear regression models were fitted for three dependent variables: normal-hours EHR usage, after-hours EHR usage, and total EHR usage. Length of hospital service was omitted as an independent variable due to multicollinearity with age. FTE status was also omitted in further analyses due to missing values. Modeling results are reported in Table 2.
Table 1. Descriptive Statistics (N=50). All values are presented as mean (SD).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Men (n=11)</th>
<th>Women (n=39)</th>
<th>Total (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>45.53 (8.20)</td>
<td>42.65 (9.95)</td>
<td>43.23 (9.69)</td>
</tr>
<tr>
<td>Length of hospital service (years)</td>
<td>12.03 (9.29)</td>
<td>8.83 (8.47)</td>
<td>9.48 (8.73)</td>
</tr>
<tr>
<td>Work relative value units</td>
<td>70.32 (51.44)</td>
<td>63.41 (40.04)</td>
<td>64.81 (42.67)</td>
</tr>
<tr>
<td>Full-time equivalent status</td>
<td>0.93 (0.09)</td>
<td>0.79 (0.19)</td>
<td>0.81 (0.18)</td>
</tr>
<tr>
<td>Clinical full-time equivalent status</td>
<td>0.56 (0.27)</td>
<td>0.53 (0.22)</td>
<td>0.54 (0.23)</td>
</tr>
<tr>
<td>Nonclinical full-time equivalent status</td>
<td>0.38 (0.25)</td>
<td>0.25 (0.20)</td>
<td>0.27 (0.22)</td>
</tr>
<tr>
<td>Normal-hours electronic health record usage (hours per week)</td>
<td>15.54 (8.91)</td>
<td>16.69 (9.16)</td>
<td>16.46 (9.12)</td>
</tr>
<tr>
<td>After-hours electronic health record usage (hours per week)</td>
<td>2.24 (3.04)</td>
<td>2.70 (3.64)</td>
<td>2.61 (3.53)</td>
</tr>
</tbody>
</table>

aThere were 13 missing values and so this variable was not used in further analyses.

Table 2. Mixed regression models.

<table>
<thead>
<tr>
<th>Model and characteristics</th>
<th>Normal-hours EHR usage</th>
<th>After-hours EHR usage</th>
<th>Total EHR usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal-hours EHR usage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(minutes)</td>
<td>_</td>
<td>–0.045 (0.010)</td>
<td>_</td>
</tr>
<tr>
<td>P value</td>
<td>_</td>
<td>&lt;.001</td>
<td>_</td>
</tr>
<tr>
<td>wRVUs</td>
<td>0.203 (0.002)</td>
<td>0.042 (0.002)</td>
<td>0.235 (0.002)</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.073 (0.0495)</td>
<td>–0.059 (0.035)</td>
<td>0.007 (0.062)</td>
</tr>
<tr>
<td>P value</td>
<td>.13</td>
<td>.08</td>
<td>.89</td>
</tr>
<tr>
<td>Gender</td>
<td>2.841 (1.234)</td>
<td>0.665 (0.907)</td>
<td>3.372 (1.580)</td>
</tr>
<tr>
<td>P value</td>
<td>.02</td>
<td>.45</td>
<td>.03</td>
</tr>
<tr>
<td>Constant</td>
<td>–2.083 (2.446)</td>
<td>2.727 (1.720)</td>
<td>0.883 (3.069)</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week</td>
<td>0.288 (0.536)</td>
<td>0.046 (0.213)</td>
<td>0.312 (0.559)</td>
</tr>
<tr>
<td>Provider</td>
<td>12.838 (3.583)</td>
<td>6.965 (2.639)</td>
<td>21.107 (4.594)</td>
</tr>
<tr>
<td>Model fitness (R^2), (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td>76.4</td>
<td>23.4</td>
<td>75.5</td>
</tr>
<tr>
<td>Random effects</td>
<td>14.4</td>
<td>52.0</td>
<td>16.2</td>
</tr>
<tr>
<td>Total</td>
<td>90.8</td>
<td>75.4</td>
<td>91.7</td>
</tr>
</tbody>
</table>

aEHR: electronic health record
bValues presented as coefficients (Standard error).
cNot applicable.
dwRVU: work relative value unit
eValues presented as Variance (Standard error).

Analyses showed a positive relationship with statistical significance between wRVUs and normal-hours EHR usage; the more wRVUs the physician generated, the more demand for normal-hours EHR usage. There was also a statistically significant relationship between gender and normal-hours EHR usage, with females spending more time using the EHRs during normal hours than males. No associations were found with age. Provider-to-provider variability (SD 3.58) contributed substantially more to the variation of normal-hours EHR usage than the variability across weeks (SD 0.53).

For after-hours EHR usage, we included normal-hours EHR usage as an explanatory variable. Modeling results show a statistically significant inverse relationship between normal hours and after-hours EHR usage; the more physicians worked with the EHRs during normal hours, the fewer hours they spent after hours with the EHRs. In addition, there was another
positive relationship with statistical significance found between wRVUs and after-hours EHR usage. The gender effect, while not statistically significant, appears to be as strong for after-hours use as it is for normal-hours use. Again, provider-to-provider variability (SD 2.64) contributed more to the variation of after-hours EHR usage than the variability across weeks (SD 0.21).

Total EHR usage was also studied using the same approach. The association with wRVU remained statistically significant and in the same direction as for the normal-hours model. Similarly, the model suggested an association between gender and total EHR usage, with female physicians spending more time than males.

Discussion

Principal Results

In this study, physicians on average spent about 16 hours weekly interacting with the EHR during their normal work hours, while spending about 3 hours weekly after hours. Thus, assuming a 2.5-day clinical work week (based on the average cumulative FTE of our sample group), our findings suggest that physicians may be spending about 6.4 hours during normal work hours and 1.2 hours after work, per weekday, completing EHR tasks. These findings are somewhat comparable to those of Arndt et al, as PCPs in their study spent 4.5 hours and 1.4 hours each weekday completing EHR tasks during and after hours, respectively [7]. Previous studies have reported time spent completing EHR tasks ranging from 2.4 to 5.9 hours per weekday; however, few of these studies specified if this time was during or after work hours [7-11].

The current study also found a positive association between wRVUs and EHR usage. This finding is expected, as wRVUs reflect the volume and intensity of medical services provided, thus the higher the wRVUs, the more likely a physician is to spend time with the EHRs. Work RVUs are one measure of physician productivity and have been found by other researchers to have a positive relationship with EHR usage [12,13]. Moreover, gender was significantly associated with normal and total EHR usage, with women spending more time with the EHRs. This association was not seen with after-hours EHR use, but perhaps while the association is the same, there is an absence of statistical significance. Nonetheless, gender differences in the use of EHRs, both during normal hours and after hours, have not been clearly documented in the scientific literature. Further, age was not found to be associated with normal or after-hours EHR use, which is somewhat consistent with findings in the scientific literature [1,14].

An unexpected finding from this study is that the provider-to-provider variation of after-hours EHR usage time was far larger than any of the other factors we examined. In terms of the variation of after-hours time spent with EHRs every week, the provider-to-provider variation explained half (52%) of the variance, whereas all fixed effects combined (wRVU, age, gender, normal-hours EHR usage) only explained 23% of the variation. Overall, the model effectively explains 75% of all the variations seen. Accordingly, one potential approach to reduce after hours use is training, which could potentially reduce the variation between providers. To our knowledge, this is the first study to quantify the effect size of this variation among providers on EHR usage.

Limitations

There are a few limitations of this study. First, the study sample size is small and limited to a single practice, a single type of provider, and a single commercial EHR system, thus limiting the generalizability of study findings. We focused our study on PCPs only, because the on-duty versus off-duty hours of hospitalists and other specialties are technically more complicated to study. Second, even though we chose to limit our study to PCPs, the use of clock time to define work hours without factoring in individual physician schedules is suboptimal. Tailoring after-hours usage to each physician’s schedule would be ideal. Third, although validated by Arndt et al, the usage hours estimated from the access logs may only be a proxy of the actual hours spent using the EHRs. Fortunately, if there is an estimation bias, it equally affects the normal-hours and after-hours usage. Fourth, some EHR activities are not patient care specific. It would be interesting to assess the differential effects of patient care–specific and nonpatient care–specific EHR activities. In addition, we did not have complete FTE status information on our sample, which further weakens the generalizability of the study findings and conclusions that can be drawn. Further, the reasons for the current findings were not empirically established. Future studies, using a mixed methods approach could shed light on reasons for these observations.

Conclusions

The present study filled a gap in the literature to statistically model variations in the duration of EHR use among PCPs and identified some of the factors that influence after-hours EHR use. These findings are essential to empirically establish these associations and advance our knowledge on this topic.

Acknowledgments

We would like to thank the following individuals, Eric Yu, Richard Hoyt, Elaine Damo and her Data Resource Center team, Dr Snyder Dane, and Macy Reese, for extracting data and helping us understand data for this study. We would also like to thank Dr John Barnard for suggesting the study idea.

Conflicts of Interest

None declared.
References


Abbreviations

EHR: electronic health record
FTE: full-time equivalent
NCH: Nationwide Children's Hospital
PCP: primary care physician
wRVU: work relative value unit
Health Care Professionals' Clinical Perspectives and Acceptance of a Blood Glucose Meter and Mobile App Featuring a Dynamic Color Range Indicator and Blood Sugar Mentor: Online Evaluation in Seven Countries

Abstract

Background: Despite many new therapies and technologies becoming available in the last decade, people with diabetes continue to struggle to achieve good glycemic control. Innovative and affordable solutions are needed to support health care professionals (HCPs) to improve patient outcomes.

Objective: To gather current self-management perceptions of HCPs in seven countries and investigate HCP satisfaction with a new glucose meter and mobile app featuring a dynamic color range indicator and a blood sugar mentor.

Methods: A total of 355 HCPs, including 142 endocrinologists (40.0%), 108 primary care physicians (30.4%), and 105 diabetes nurses (29.6%), were recruited from the United Kingdom (n=50), France (n=50), Germany (n=50), India (n=54), Algeria (n=50), Canada (n=51), and the United States (n=50). HCPs experienced the OneTouch Verio Reflect glucose meter and the OneTouch Reveal mobile app online from their own office computers using interactive demonstrations via webpages and multiple animations. After providing demographic and clinical practice insights, HCPs responded to statements about the utility of the system.

Results: Concerning current practice, 83.1% (295/355) of HCPs agreed that poor numeracy or health literacy was a barrier for their patients. A total of 85.9% (305/355) and 92.1% (327/355) of HCPs responded that type 2 diabetes (T2D) and type 1 diabetes (T1D) patients were aware of what represented a low, in-range, or high blood glucose result. Only 62.0% (220/355) felt current glucose meters made it easy for patients to understand if results were in range. A total of 50.1% (178/355) and 78.0% (277/355) of HCPs were confident that T1D and T2D patients took action for low or high results. A total of 87.0% (309/355) agreed that the ColorSure Dynamic Range Indicator could help them teach patients how to interpret results and 88.7% (315/355) agreed it made them more aware of hyper- and hypoglycemic results so they could take action. A total of 83.7% (297/355) of HCPs agreed that the Blood Sugar Mentor feature gave personalized guidance, insight, and encouragement so patients could take action. A total of 82.8% (294/355) of HCPs also agreed that the Blood Sugar Mentor provided real-time guidance to reinforce the goals HCPs had set so patients could take steps to manage their diabetes between office visits. After experiencing the full system, 85.9% (305/355) of HCPs agreed it was beneficial for patients with lower numeracy or health literacy; 96.1% (341/355) agreed that it helped patients understand when results were low, in range, or high; and 91.0% (323/355) agreed that the way it displayed diabetes information would make patients more inclined to act upon results. A total of 89.0% (316/355) of HCPs agreed that it would be helpful for agreeing upon appropriate in-range goals for their patients for their next clinic visit.

Conclusions: This multi-country online study provides evidence that HCPs were highly satisfied with the OneTouch Verio Reflect meter and the OneTouch Reveal mobile app. Each of these use color-coded information and the Blood Sugar Mentor...
feature to assist patients with interpreting, analyzing, and acting upon their blood glucose results, which is particularly beneficial to keep patients on track between scheduled office visits.

(JMIR Hum Factors 2019;6(3):e13847) doi:10.2196/13847

KEYWORDS
ColorSure Dynamic Range Indicator; Blood Sugar Mentor; mobile app; blood glucose meter; self-monitoring of blood glucose; health care professionals

Introduction

Despite many new technologies and medications to treat diabetes during the past decade, the proportion of patients in the United States achieving a glycated hemoglobin (HbA1c) target of less than 7.0% has been fairly constant from 2003 to 2010, remaining at just over 50% [1]. A further analysis confirmed that those achieving an HbA1c of less than 7.0% only slightly declined from 52.2% to 50.9% by 2014, and the proportion in poor control (ie, HbA1c >9.0%) actually worsened from 12.6% to 15.5% from 2007 to 2014 [2]. Even among patients with type 1 diabetes (T1D) who are part of the so-called “T1D exchange” dataset in specialty diabetes centers in the United States, mean HbA1c worsened from 7.8% in 2010–2012 to 8.4% in 2016–2018, despite continuous glucose monitoring usage increasing from 7% to 30% in this population [3]. These findings are concerning and would seem to indicate that other factors are preventing patients from taking full advantage of self-monitoring or continuous glucose monitoring technologies, which may include patients simply struggling to interpret and act upon information from these technologies. In fact, the latest American Diabetes Association guidelines suggest that optimal use of self-monitoring technologies requires proper review and interpretation of the data by both the patient and the provider to ensure that data are used in an effective and timely manner; the guidelines also suggest that patients should be taught how to use their self-monitoring data to adjust food intake, exercise, or medications to achieve specific goals [4]. Appropriate education addressing how to interpret self-monitoring of blood glucose information and how to respond to out-of-range results have been identified as important requirements for useful self-monitoring of blood glucose practice [5]. However, a survey of 886 people showed that approximately 50% of insulin and non-insulin-using patients with type 2 diabetes (T2D) regularly took no action for out-of-range readings, low or high, with any self-care adjustments [6]. Disparities in literacy and low numeracy in patients in various countries could impede efforts to support patients who struggle to comprehend self-care guidance or use the self-monitoring technologies provided by health care professionals (HCPs). For example, low diabetes-related numeracy skills are associated with fewer self-management behaviors [7], and poor numeracy is also associated with suboptimal glycemic control in patients with T2D [8] and T1D [9]. We previously reported that glucose meters utilizing color range indicators improved the ability of patients to interpret glucose results [10], improved decision making in terms of taking action [11], and improved glycemic control when patients switched from other glucose meters [12-14]. In this study, we solicited feedback from HCPs in seven countries to explore current clinical practice issues with respect to self-monitoring and to determine how the OneTouch Verio Reflect meter, as well as the OneTouch Reveal mobile app, with ColorSure Dynamic Range Indicator and Blood Sugar Mentor (BSM) features might provide benefits for patients in these countries.

Methods

Materials

OneTouch Verio Reflect (LifeScan) is a blood glucose meter intended for self-testing by people with diabetes. The meter automatically analyzes high and low glucose patterns, tracks trends, and provides on-meter guidance messages to help patients understand and manage their glucose levels and help them detect excursions above or below a desired glucose range. The meter has a ColorSure Dynamic Range Indicator to show when a result is low (blue), in range (green), or high (red) using seven color-coded segments. The patient or HCP can personalize the ranges of this feature and choose between an arrow or emoji character to indicate which colored segment relates to the current reading. The meter has additional features over basic glucose meters, including a test goal tracker; advanced tagging of results with various event symbols, such as carbohydrates, stress, illness medication, or exercise; a color grid of the last 30 days of results (ie, time of day/glucose range); and provides an on-meter trend graph of average glucose values (see Figure 1).

Furthermore, the on-meter Blood Sugar Mentor feature provides a variety of mentor tips, pattern messages, and awards when certain testing achievements or glycemic goals are met (see Figure 2), such as the following:

1. Mentor Tips are provided when results are consistently in range or are currently trending low or high.
2. Pattern Messages are provided when the meter identifies a pattern of glucose results that fall outside the high and low range limits the user sets in the meter.
3. Awards are earned when award criteria are met, such as meeting the daily test goal.

The Verio Reflect meter has Bluetooth capability to allow transfer of blood glucose readings and event tagging information to the OneTouch Reveal app (eg, on a mobile phone or tablet), which has additional insight and trending features and provides HCPs and patients with the ability to share information during and between clinic visits (see Figure 3).
Figure 1. OneTouch Verio Reflect meter showing a selection of screens.

Figure 2. Selected Blood Sugar Mentor screens on the OneTouch Verio Reflect meter.
Figure 3. Selected Blood Sugar Mentor messages on the OneTouch Reveal app.

Procedure
This multi-country online survey study was conducted by individual HCPs from institutions and clinical practices within each country. Webpages containing text information, product and feature images, and links to short animations describing meter and app features were provided to the HCPs. A total of 355 HCPs from seven countries—the United Kingdom, France, Germany, India, Algeria, Canada, and the United States—were recruited and included endocrinologists, primary care physicians (PCPs), and diabetes nurses. The sampling strategy dictated a target of 20 endocrinologists, 15 physicians, and 15 diabetes nurses per country to ensure we had representative views from different HCPs. Online inclusion criteria questions precluded HCPs who did not routinely treat at least 10 T1D or T2D patients per week or who did not have at least 15% of their patient population currently on insulin therapy. Before the online experience, all HCPs provided demographic and clinical practice metrics with respect to the patients they routinely advised or treated.

HCPs were asked eight clinical practice questions to determine the confidence they had in the ability of their current patients with type 1 and type 2 diabetes to interpret or act upon blood glucose data. Participating HCPs were then presented with an online experience sharing identical capability, functionality, and navigation as the intended meter and app products. The meter and app animations were preloaded with representative glucose results, messages, guidance, or information that provided examples of the meter screens that would appear whenever HCPs or patients review information. The HCPs interacted online with a series of webpages displaying both text and visuals of the meter and app, with embedded links on the selected webpages allowing HCPs to watch short product animations. At various points during these activities, 24 survey questions were presented to assess the HCPs’ opinions of the value of various functions and features of the meter and app. After completing online activities, HCPs were asked four clinical practice-based questions pertaining to the value of the meter and app for supporting patients in future with diabetes self-management.

Statistical Analyses
Continuous demographic variables were described as median and range or mean and standard deviation. Categorical demographic variables were described as percentages within categories and are presented with both numerators and denominators. HCP responses to survey statements were recorded using a 5-point Likert scale with a favorable response (4 or 5) being met if the lower 95% one-sided confidence limit for the percentage of participants providing a favorable response per item was greater than 50%.

Results
Health Care Professionals’ Demographic and Clinical Practice Information
A total of 355 HCPs took part in the study from seven countries: the United Kingdom (n=50), France (n=50), Germany (n=50), India (n=54), Algeria (n=50), Canada (n=51), and the United States (n=50). Out of the 355 HCPs, professional backgrounds included 142 endocrinologists (40.0%), 108 PCPs (30.4%), and 105 diabetes nurses (29.6%). Mean age across all seven countries was 44.8 years (SD=9.6). The majority were male (n=233, 65.6%), had more than 10 years of experience (n=230, 65.0%), and had been in practice for more than 10 years (n=226, 63.6%). The majority of HCPs reported treating at least 10 T1D patients per week (n=308, 86.7%) and at least 10 T2D patients per week (n=304, 85.7%). The majority of patients were on insulin therapy (n=338, 95.2%), and the majority of patients were on insulin pumps (n=332, 93.5%). The majority of patients were on multiple daily insulin injections (n=339, 95.3%).
countries was 48 years (SD 9) for endocrinologists, 49 years (SD 9) for PCPs, and 46 years (SD 9) for diabetes nurses. The proportion of patients with T1D and T2D, respectively, typically seen by each professional in routine clinical practice was 25% and 69% for endocrinologists, 16% and 80% for PCPs, and 27% and 67% for diabetes nurses (see Table 1).

Table 1. Health care professionals’ status and patient population information.

<table>
<thead>
<tr>
<th>Patient population information</th>
<th>Health care professionals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endocrinologists (n=142)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>99 (69.7)</td>
</tr>
<tr>
<td>Female</td>
<td>43 (30.3)</td>
</tr>
<tr>
<td>Age in years, mean (SD)</td>
<td>48 (9)</td>
</tr>
</tbody>
</table>
| Patient diabetes type, mean %
|                              |                           |                               |                          |                 |
| Type 1 diabetes               | 25                        | 16                            | 27                       | 23              |
| Type 2 diabetes               | 69                        | 80                            | 67                       | 72              |
| Other                         | 6                         | 4                             | 6                        | 5               |
| Patient-specific therapy, mean%
|                              |                           |                               |                          |                 |
| Diet and exercise             | 9                         | 12                            | 11                       | 11              |
| Diabetes medications only     | 31                        | 45                            | 29                       | 35              |
| Diabetes medications and insulin | 28                      | 24                            | 27                       | 26              |
| Insulin only (injections or pump) | 22                   | 14                            | 23                       | 20              |
| Insulin plus CGM\textsuperscript{a} or FGM\textsuperscript{b} | 13                        | 9                             | 14                       | 12              |

\textsuperscript{a}Percentages shown are estimates given by the health care professionals.

\textsuperscript{b}CGM: continuous glucose monitor.

\textsuperscript{c}FGM: flash glucose monitor.

Health Care Professionals’ Clinical Practice Feedback on Patient Self-Care

Of the HCPs who responded, 83.1% (295/355) confirmed that poor numeracy or health literacy was a barrier for many of their patients to improve their diabetes care. However, a high percentage of HCPs responded that their current patients were aware about what represents a low, in-range, or high glucose result when it appeared on their glucose meter, with 85.9% (305/355) in agreement for T2D patients and 92.1% (327/355) for T1D patients. Far fewer HCPs were convinced that the blood glucose meters used by their patients make it easy for them to understand if their results are in range, with only 62.0% (220/355) in agreement. In terms of confidence that their patients understood the reasons why they sometimes got low or high glucose results, the responses from HCPs were quite different in that only 49.0% (174/355) agreed that patients with T2D understood the reasons why, with a higher percentage of HCPs (263/355, 74.1%) agreeing that their T1D patients understood the reasons for low or high glucose results. This disparity was also evident when HCPs were asked how confident they were that their patients took action when they got a low or high glucose result, with only 50.1% (178/355) agreeing that patients with T2D took action, with a higher percentage of HCPs (277/355, 78.0%) agreeing that their T1D patients would take action for low or high glucose results (see Figure 4).
Health Care Professionals’ Feedback During Online Experiences With the Meter and App

During the interactive online meter experience, 91.5% (130/142) of endocrinologists, 91.7% (99/108) of PCPs, and 89.5% (94/105) of nurses agreed that the ColorSure Dynamic Range Indicator could help patients understand when glucose results were near the high or low limits, so they could make adjustments before going out of range. A total of 86.6% (123/142) of endocrinologists, 88.9% (96/108) of PCPs, and 91.4% (96/105) of nurses also agreed that the ColorSure Dynamic Range Indicator could help patients be more aware of hyper- and hypoglycemic results so they could take action (see Table 2).
<table>
<thead>
<tr>
<th>Survey statement</th>
<th>Favorable responses by health care professionals, n (%)</th>
<th>Combined (N=355)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endocrinologists (n=142)</td>
<td>PCPs (n=108)</td>
</tr>
<tr>
<td>The ColorSure Dynamic Range Indicator can help patients understand when glucose</td>
<td>130 (91.5)</td>
<td>99 (91.7)</td>
</tr>
<tr>
<td>results are near high or low, so they can make adjustments before going out of  range.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The ColorSure Dynamic Range Indicator will help patients be more aware of hyper</td>
<td>123 (86.6)</td>
<td>96 (88.9)</td>
</tr>
<tr>
<td>and hypo results so they can take action.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe the meter will help patients reduce the number of hypoglycemic episodes</td>
<td>106 (74.6)</td>
<td>86 (79.6)</td>
</tr>
<tr>
<td>they experience.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients can know if their actions are working with the enhanced blue, green, red</td>
<td>125 (88.0)</td>
<td>99 (91.7)</td>
</tr>
<tr>
<td>ColorSure Dynamic Range Indicator that shows if they’re in range, out of range,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>or near a high or low, and see that information directly on their smartphone.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The solution’s ColorSure technology could make it easy to teach my patients to</td>
<td>122 (85.9)</td>
<td>94 (87.0)</td>
</tr>
<tr>
<td>interpret their blood glucose results.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Blood Sugar Mentor will help patients understand the impact of food, activity,</td>
<td>122 (85.9)</td>
<td>95 (88.0)</td>
</tr>
<tr>
<td>and medication on their glucose, so they can make adjustments to improve blood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sugar control.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Blood Sugar Mentor automatically identifies times when patients are likely to</td>
<td>123 (86.6)</td>
<td>88 (81.5)</td>
</tr>
<tr>
<td>experience hyper and hypo events and alerts them, so they will be able to make</td>
<td></td>
<td></td>
</tr>
<tr>
<td>changes to their daily routine.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Blood Sugar Mentor analyses patterns, tracks trends, and could help guide</td>
<td>122 (85.9)</td>
<td>95 (88.0)</td>
</tr>
<tr>
<td>patients toward better self-management and staying healthy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The guidance, insights, and encouragement provided by the Blood Sugar Mentor will</td>
<td>115 (81.0)</td>
<td>94 (87.0)</td>
</tr>
<tr>
<td>help patients take actions to manage their diabetes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The real-time guidance that the Blood Sugar Mentor provides will help reinforce</td>
<td>115 (81.0)</td>
<td>90 (83.3)</td>
</tr>
<tr>
<td>the goals you set, so patients can take steps to manage their diabetes between</td>
<td></td>
<td></td>
</tr>
<tr>
<td>office visits.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe the meter will provide patients with greater understanding and guidance,</td>
<td>117 (82.4)</td>
<td>96 (88.9)</td>
</tr>
<tr>
<td>so they can confidently make progress managing their blood sugar.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe the meter will help patients stay on top of their testing routine and</td>
<td>106 (74.6)</td>
<td>94 (87.0)</td>
</tr>
<tr>
<td>control their glucose around meals, activities, and specific times of day.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe the meter will help patients manage their blood sugar more effectively</td>
<td>109 (76.8)</td>
<td>93 (86.1)</td>
</tr>
<tr>
<td>than devices without a ColorSure Dynamic Range Indicator and Blood Sugar Mentor.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The solution’s Blood Sugar Mentor will make it easy for patients to see and</td>
<td>124 (87.3)</td>
<td>97 (89.8)</td>
</tr>
<tr>
<td>understand how their lifestyle choices impact their blood glucose levels, right</td>
<td></td>
<td></td>
</tr>
<tr>
<td>on their meter and smartphone.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With the Blood Sugar Mentor, patients get personalized guidance, insight, and</td>
<td>122 (85.9)</td>
<td>95 (88.0)</td>
</tr>
<tr>
<td>encouragement so they can take actions based on their current and previous results.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The solution’s Blood Sugar Mentor could help patients be more proactive in</td>
<td>116 (81.7)</td>
<td>88 (81.5)</td>
</tr>
<tr>
<td>managing their glucose levels.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The ongoing guidance provided by the solution’s Blood Sugar Mentor could help my</td>
<td>115 (81.0)</td>
<td>97 (89.8)</td>
</tr>
<tr>
<td>struggling patients improve their understanding of diabetes management and take</td>
<td></td>
<td></td>
</tr>
<tr>
<td>action.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With this solution, patients do not have to memorize facts, numbers, and actions,</td>
<td>114 (80.3)</td>
<td>90 (83.3)</td>
</tr>
<tr>
<td>so they can confidently manage their diabetes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The ColorSure Dynamic Range Indicator helps patients know when they are near</td>
<td>119 (83.8)</td>
<td>90 (83.3)</td>
</tr>
<tr>
<td>hypoglycemic levels and provides them with actions they could take to help avoid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>them.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The solution will make it easy to quickly see and assess my patients’ lifestyle</td>
<td>103 (72.5)</td>
<td>91 (84.3)</td>
</tr>
<tr>
<td>and blood glucose data and help me to provide more personalized care.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The solution alerts patients when they are near hyper- or hypoglycemic levels and</td>
<td>126 (88.7)</td>
<td>89 (82.4)</td>
</tr>
<tr>
<td>provides simple suggestions for corrective actions to help avoid them.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>This solution provides patients with greater understanding and guidance in</td>
<td>121 (85.2)</td>
<td>95 (88.0)</td>
</tr>
<tr>
<td>managing their blood sugar so they can confidently make progress toward their</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diabetes management goals.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Minimizing the incidence of hypoglycemic events is a key part of diabetes care and a real concern for patients and HCPs. A total of 74.6% (106/142) of endocrinologists, 79.6% (86/108) of PCPs, and 73.3% (77/105) of nurses believed that this new meter will help patients reduce the number of hypoglycemic episodes they experience. This is an important endorsement and part of this advocacy may relate to the HCPs’ experience of the BSM features on the meter and app. A total of 86.6% (123/142) of endocrinologists, 81.5% (88/108) of PCPs, and 89.5% (94/105) of nurses agreed that the BSM automatically identified times when patients are likely to experience hyper- and hypoglycemic events and alerts them so they will be able to make changes to their daily routine. A total of 81.0% (115/142) of endocrinologists, 87.0% (94/108) of PCPs, and 83.8% (88/105) of nurses agreed that the guidance, insights, and encouragement provided by the BSM would also help their patients take action to manage their diabetes. Increasingly, it is what patients do themselves between visits that concerns HCPs. A total of 81.0% (115/142) of endocrinologists, 83.3% (90/108) of PCPs, and 84.8% (89/105) of nurses agreed that the real-time guidance that the BSM provides will help reinforce the goals that they set so patients can take steps to manage their diabetes between office visits. Overall, combining the attributes of the meter and the app together (ie, the solution), 88.7% (126/142) of endocrinologists, 92.6% (100/108) of PCPs, and 91.4% (96/105) of nurses agreed that the solution alerts patients when they are near hyper- or hypoglycemic levels and provides simple suggestions for corrective actions to help avoid them. In addition, 85.2% (121/142) of endocrinologists, 88.0% (95/108) of PCPs, and 85.7% (90/105) of nurses agreed that the solution provides patients with greater understanding and guidance in managing their blood sugar so they can confidently make progress toward their diabetes management goals.

### Health Care Professionals’ Clinical Practice Outlook for Patients Based on Meter and App Experiences

Having seen the new meter and app, 87.3% (124/142) of endocrinologists, 80.6% (87/108) of PCPs, and 88.6% (93/105) of nurses agreed it was beneficial for helping patients with lower numeracy or health literacy who may struggle to interpret results. In terms of supporting comprehension of glucose data, 95.1% (135/142) of endocrinologists, 94.4% (102/108) of PCPs, and 98.1% (103/105) of nurses agreed the system was beneficial for helping patients understand when their glucose results are low, in range, or high; 87.3% (124/142) of endocrinologists, 92.6% (100/108) of PCPs, and 95.2% (100/105) of nurses agreed that displaying diabetes information in this way would make their patients more inclined to act upon low or high results. With respect to improving patient consultations, 87.3% (124/142) of endocrinologists, 88.9% (96/108) of PCPs, and 92.4% (97/105) of nurses agreed this new system would be a helpful tool during patient conversations to agree upon appropriate in-range goals for their patients’ next clinic visit (see Figure 5).

---

**Survey statement**

| Favorable responses by health care professionals, n (%)<sup>a</sup> |
|---|---|---|---|---|
| **Endocrinologists** (n=142) | **PCPs**<sup>b</sup> (n=108) | **Nurses** (n=105) | **Combined** (N=355) |
| The solution analyses and informs patients of blood glucose trends and potential causes, and provides personalized guidance patients can use to avoid hyper- and hypoglycemia. | 120 (84.5) | 93 (86.1) | 92 (87.6) | 305 (85.9) |
| The solution gives a better understanding of the underlying causes of patients’ blood glucose changes so they are better able to manage them. | 112 (78.9) | 90 (83.3) | 81 (77.1) | 283 (79.7) |

<sup>a</sup>Results shown are favorable responses (strongly agree or agree) on a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree). All favorable responses met the acceptance criteria (ie, lower bound of 95% confidence limits >50%).

<sup>b</sup>PCP: primary care physician.
Figure 5. Health care professionals' (HCPs) clinical practice outlook after experiencing the OneTouch Verio Reflect meter and the OneTouch Reveal mobile app. Results shown are percentage favorable responses (strongly agree or agree; very beneficial or beneficial) on a 5-point Likert scale ranging from 1 (strongly agree or very beneficial) to 5 (strongly disagree or not at all beneficial). All percentage favorable responses met the acceptance criteria (ie, lower bound of 95% confidence limits >50%).

**Discussion**

This online study suggests that HCPs from seven countries had high overall acceptance of a new glucose meter and mobile app. This study also confirmed that using color-coded information and a BSM feature to assist patients with interpreting and acting upon their blood glucose information is an approach that resonates universally with different types of HCPs from different health care environments. We gathered feedback from a large sample of HCPs from each of seven countries and ensured that, collectively across all seven countries, we had a large number of endocrinologists, PCPs, and diabetes nurses to enable robust survey responses.

We found good agreement between the three types of HCPs in terms of the current self-care practices of patients and the issues that they face. HCPs were broadly very concerned that poor numeracy or health literacy were stumbling blocks that prevented their patients from being successful in aspects of diabetes care. Furthermore, it was interesting that HCPs felt that the majority of their patients had a good awareness about what represents a low, in-range, or high result; this viewpoint was expressed on behalf of patients with T1D and, to a marginally lesser extent, on behalf of patients with T2D. Surprisingly, given a more direct question specifically asking if the current glucose meters used by their patients made it easy for them to understand if their results were in range, agreement from HCPs of all types plummeted to between 55% and 65%. It is somewhat disappointing that after almost 30 years of companies developing home blood glucose meters, only around 60% of HCPs feel these devices provide this valuable context to patients, in parallel to their numerical result. Immediately alerting the patient to whether their on-screen result is in range, low, or high should be an expectation from the latest blood glucose meters. Interestingly, HCPs were skeptical that most of their T2D patients understood the reasons why they got low or high results, with less than 50% of HCPs agreeing that they did. Conversely, over 70% of HCPs agreed that T1D patients understood why they got low or high results, which is a higher percentage, although not a ringing endorsement and may be expected given that T1D patients test more frequently and routinely make therapy decisions based on results. In contrast to the HCP feedback in this study where between 86% and 96% of HCPs felt that their patients were aware of the context of their glucose results, our previous research with patients contradicts this feedback. We found that T1D and T2D patients struggle to categorize glucose results as low, in-range, or high but did improve after experiencing a color range indicator feature explaining how to categorize different glucose readings [10]. Furthermore, our study showed that only 50% of HCPs felt T2D patients took action for low or high results; this low percentage mirrors the feedback obtained in a survey of insulin-using T2D patients in the United States who also responded that they would only take action for around 50% of their low or high glucose values [6]. Unsurprisingly, HCPs were more optimistic that T1D patients would take action for lows or highs, perhaps responding with more urgency based on the advice from the HCPs to avoid prolonged exposure to hypo- or hyperglycemia.
Technological advances now and in the future will present both opportunities and challenges. There is a real concern that HCPs may become overloaded with patient data from multiple devices (ie, continuous glucose monitors [CGMs], insulin pumps, health records, activity trackers, and health monitors), with an expectation that they analyze and translate copious amounts of information into actions for their patients [15]. Given that it remains the case that the average face-to-face time spent by a physician with a patient with diabetes is measured in mere minutes [16], it is essential to provide solutions that offer simple, automatic, data-driven advice directly to patients between consultations. Clearly, CGM companies are seeking to provide greater insight and decision support to patients between visits to improve glycemic control, although it must be acknowledged that these products remain relatively expensive and, for many patients, it may be worth first transitioning to more-advanced glucose meters. In this regard, HCPs in our online study experienced a BSM feature that resides directly on the meter and can communicate via Bluetooth to the Reveal mobile app, which provides additional BSM features. Our HCPs felt that the automatic insights and analysis of the glucose data provided by the BSM had important benefits in comparison to systems they use now, including automatically identifying times when patients are likely to experience hyper- and hypoglycemic events, analyzing patterns and tracking trends to help guide patients toward better self-management, and giving insights and encouragement to help patients take action. These are all areas to support patients between consultations to enable them to stick to the goals that have been set with HCPs during those infrequent office visits. Given the propensity for depression and anxiety among some diabetes patients, it was particularly important to ensure that the pattern and BSM messages conceived on our system were positive and routinely focused on getting back in range or achieving more in-range results rather than demotivating patients.

The most encouraging aspect of the surveys was how HCPs responded to statements about how the features and benefits of this new meter and app might contribute to the future care of their diverse patient populations. The diabetes community has been exploring new end points for diabetes care, going beyond HbA1c, and the idea of focusing on time or data-in-range has gained traction [17-19]. HCPs in our study agreed that our new meter and app would help them set appropriate in-range goals between visits, given that the meter clearly indicates, by pointing to a green bar, when results are in range. Facilitating this new and very practical way of ensuring patients maintain glycemic control is a key benefit. HCPs also acknowledged, prestudy, that they are concerned about basic numeracy and literacy and how it impacts their patients’ ability to self-manage. It was evident that the majority of HCPs felt the new meter and app could overcome some of these challenges to allow patients to better interpret and act upon their blood glucose data. It is worth noting some wider variations between HCPs for certain questions. For example, fewer endocrinologists (75%) agreed that the meter would help patients stay on top of their testing routine and control their glucose around meals, activities, and specific times of day, compared to 87% of PCPs and 82% of nurses. Similarly, fewer endocrinologists (73%) agreed the meter and app would allow them to quickly see and assess their patients’ lifestyle and blood glucose data, compared to 84% of PCPs and 86% of nurses. On the whole, PCPs and nurses were more aligned on their responses than endocrinologists, which may be partly explained by the more complex patients referred for secondary or specialty care.

In terms of study limitations, it is worth noting that these results may not be generalizable to HCPs in other countries that were not part of the survey and that despite recruiting a good sample of at least 50 HCPs per country, this is sample is only representative of the views of HCPs in each country. The plethora of color-enhanced features and displays now becoming routinely available from most manufacturers on glucose meters, diabetes apps, and sensor technologies (ie, CGMs) should be considered by HCPs with respect to those patients who may have some form of color visual impairment.

In conclusion, this multi-country online study provides evidence that HCPs were highly satisfied with the OneTouch Verio Reflect meter and OneTouch Reveal mobile app. Each of these use color-coded information and a BSM feature to assist patients with interpreting, analyzing, and acting upon their blood glucose results, which is particularly beneficial to keep patients on track during and between scheduled office visits.

Acknowledgments
The authors thank Ipsos, Parsippany, NJ, for the creation of webpages and HCP recruitment. This study was funded by LifeScan, Inc.

Conflicts of Interest
MG is an employee of LifeScan Scotland. UV, KR, and GH are all employees of LifeScan Global Corporation. OS is a member of Forschergruppe Diabetes, Neuherberg, Germany. OS received some financial support for his contribution and review of the article.

References


Abbreviations

BSM: Blood Sugar Mentor
CGM: continuous glucose monitor
FGM: flash glucose monitor
HbA_1c: glycated hemoglobin
HCP: health care professional
PCP: primary care physician
T1D: type 1 diabetes
T2D: type 2 diabetes
Tailoring of a Smartphone Smoking Cessation App (Kick.it) for Serious Mental Illness Populations: Qualitative Study

Pauline Klein1,2*, MPH; Sharon Lawn1*, PhD; George Tsourtos2*, PhD; Joep van Agteren3, MSc

1Flinders Human Behaviour and Health Research Unit, Department of Psychiatry, Flinders University, Adelaide, Australia
2College of Medicine and Public Health, Flinders University, Adelaide, Australia
3Wellbeing and Resilience Centre, South Australian Health and Medical Research Institute, Adelaide, Australia
*these authors contributed equally

Corresponding Author:
Pauline Klein, MPH
College of Medicine and Public Health
Flinders University
GPO Box 2100
Adelaide, 5001
Australia
Phone: 61 427 370 673
Email: klei0081@flinders.edu.au

Abstract

Background: Smoking rates of Australians with severe mental illness (SMI) are disproportionately higher than the general population. Despite the rapid growth in mobile health (mHealth) apps, limited evidence exists to inform their design for SMI populations.

Objective: This study aimed to explore the feasibility, acceptability, and utility of adapting a novel smoking cessation app (Kick.it) to assist smokers with SMI to prevent smoking relapse and quit.

Methods: Using co-design, two in-depth interviews with 12 adult smokers and ex-smokers with SMI were conducted in this qualitative study. Stage 1 interviews explored participants’ smoking-related experiences and perceptions of social support for smoking cessation, informed the development of the stage 2 interview schedule, and provided context for participants’ responses to the second interview. Stage 2 interviews explored participants’ perceptions of the feasibility, utility, and acceptability of the app features for SMI populations.

Results: People with SMI perceived mHealth interventions to support their quit smoking attempts as feasible, acceptable, and useful. Key emerging themes included personalization of the app to users’ psychosocial needs, a caring app to mediate self-esteem and self-efficacy, an app that normalizes smoking relapse and multiple quit attempts, a strong focus on user experience to improve usability, and a social network to enhance social support for smoking cessation.

Conclusions: This study gained an in-depth understanding of the lived experiences of smoking and quitting among people with SMI and their perception of the Kick.it app features to help inform the tailoring of the app. Specific program tailoring is required to assist them in navigating the complex interactions between mental illness and smoking in relation to their psychosocial well-being and capacity to quit. This study describes the adaptations required for the Kick.it app to meet the specific needs and preferences of people with SMI. Results of this study will guide the tailoring of the Kick.it app for SMI populations. The study findings can also inform a co-design process for the future development and design of smoking cessation apps for SMI populations.

(JMIR Hum Factors 2019;6(3):e14023) doi:10.2196/14023

KEYWORDS
mental health; mHealth; tobacco; smoking cessation; public health; technology

https://humanfactors.jmir.org/2019/3/e14023
**Introduction**

**Background**

Tobacco smoking is a major cause of preventable mortality and morbidity [1] and health inequalities for people with severe mental illness (SMI), including extreme social, economic, and physical health disadvantages [2]. SMI diagnoses include, for example, schizophrenia and bipolar disorder. In Australia, the smoking rates for SMI populations have remained inequitably high [3] for more than a decade [4]. Smoking rates for SMI populations in South Australia are alarmingly high, at almost triple (43.6%), that of the general population (15%) [5]. The high prevalence of smoking in this population is indicative of mental illness and smoking being intrinsically linked [4,6]. This is evident as people with mental illness often identify as smokers [6] and use smoking as a form of self-medication to help them cope with and relieve their symptoms of mental illness [4,6,7]. Owing to the inseparable nature of mental illness and smoking, nicotine addiction in this population has been difficult to treat [4].

Smoking-related studies have revealed that most people with SMI want to quit [8] and often attempt to quit but can find it challenging to quit without support [4,9]. A review of smoking cessation interventions, such as motivational interviewing, found limited evidence to support the interventions’ effectiveness in assisting people with schizophrenia to quit [10]. Cutting-edge digital health technology, such as mobile health (mHealth) smoking cessation apps [11-13], may contribute to the solutions needed to address this significant public health problem [11,14].

There are currently hundreds of smoking cessation apps available for download; however, limited studies have been conducted to assess the quality of app design. Research assessing the quality of generic smoking cessation apps revealed that most do not adhere to best-practice guidelines for smoking cessation, such as recommending pharmacotherapy [15,16]. Many smoking cessation apps also rated low on technical quality [17]. A review of 112 smoking cessation apps found that only 6 of these apps rated high on technical quality, such as having aesthetic appeal [16]. Despite the lack of quality for the vast majority of apps, scientific studies on smartphone interventions are promising in increasing cessation. Smokers who received the smoking cessation interventions demonstrated as much as a 1.7 times higher quit rate than smokers who did not receive the interventions [18].

In contrast to the vast availability of generic smoking cessation apps, there are only 3 that have been tailored for SMI populations [19-21]. This highlights a substantial gap in the availability of smoking cessation apps for SMI populations. This is particularly important when considering the cognitive impairments many people with SMI have to endure [22], which limits their ability to use apps [15,23,24]. A study on the QuitPal app developed by the National Cancer Institute, found that people with SMI experienced problems navigating the app, such as entering data, which was particularly relevant among participants with cognitive impairments and tremors [25]. Vilardaga et al used co-design, a person-centered approach to technology design, to involve people with SMI in the tailoring of the Learn to Quit app, developed by the University of Washington. Findings indicated that adapting a user experience (UX) approach with simple functionality, including large buttons and simple screens, improved the utility, usability, and acceptability of the app among people with SMI [21]. UX is a human-centered approach to improving end users’ performance and their psychological experience of technology systems [26]. Furthermore, exploration of participants’ preferences for app features found that people with SMI were interested in gamification (application of game design components and game principles in nongame systems) [27], interactive strategies to develop quit skills, and tracking devices for monetary incentives [21,25].

Building on this existing knowledge, there is a need to gain a deeper understanding of the relationship between mental illness, smoking, and smoking cessation to inform the design of effective smoking cessation app approaches for this population. This can help to guide how an app can be tailored to meet the specific requirements of people with SMI to reduce their smoking and quit. Currently, there are no smoking cessation apps for SMI populations that have been investigated within an Australian context, regarding their acceptability, feasibility, and usefulness. This study aims to address these important issues in relation to tailoring the Kick.it app for SMI populations [14].

**The Kick.it App and Its Theoretical Frameworks**

Kick.it is a generic Australian-based prototype app that was originally co-designed for use by the general population of smokers [28], using intervention mapping (IM), which is a rigorous multitheoretical intervention development framework [29]. This consisted of a comprehensive needs analysis of the literature and stakeholder input from health professionals and smokers to identify the problem behaviors and determinants for smoking cessation. A co-design principle has also been used in this study to tailor the app for SMI populations before releasing the app on the marketplace. The design of the app for SMI populations enables app users to create a profile (ie, input information about their psychiatric diagnoses and smoking) and receive a personalized quit program that offers smoking cessation approaches tailored to meet their unique needs. These smoking cessation approaches are based on multitheoretical perspectives [28], as follows.

The Theoretical Domains Framework, a valid multitheoretical approach [30], underpinned the determinants for smoking cessation (eg, knowledge and skills) and the change objectives required to assist app users to quit (eg, increased knowledge of and ability to implement quit strategies) [28]. The Behavior Change Technique Taxonomy (v1) was used to identify behavior change approaches [31] and behavior change outcomes. These were then translated into app features and practical applications for smoking cessation [28]. The Persuasive System Design, a framework for technology development that targets attitude and behavior change, was also applied to inform the choice of app features [32].

There are 4 core features contained in the Kick.it app. The smoke and crave profile feature is based on the principles of ecological momentary assessment [33], which tracks user’s smoking and quitting behaviors in real time and delivers in-time quit strategies.
at critical moments to prevent smoking relapse and support smoking cessation (see Multimedia Appendices 1 and 2 for screenshots of in-time interventions). When an app user logs a smoke or crave, it activates the tracking device and provides them with a progress report [28] (see Multimedia Appendix 3 for screenshot of the tracking device). The Kick.it stack feature contains education and strategies to assist app users during their quit attempt (see Multimedia Appendix 4 for screenshot of an educational video on nicotine replacement therapy). The social network platform is a unique app feature that leverages peer support and normative social influence for smoking cessation through chatrooms, community feeds, and links to other app users’ social networks [28]. To our knowledge, Kick.it is the first app to include a social network feature to enhance social support for SMI populations. The missions and treatment goals feature is based on an incentive and reward system that encourages app users to engage in daily health-enhancing activities and log their pharmacotherapy use. A comprehensive overview of the development and design of the Kick.it app has been published elsewhere [28]. Figure 1 presents the Kick.it app features and practical applications [28].

Assessing the feasibility, utility, and acceptability of tailoring the Kick.it app for SMI populations provides an ideal opportunity to address the limited availability of smoking cessation apps for this population.

**Figure 1.** Kick.it app features and practical applications.

<table>
<thead>
<tr>
<th>Features</th>
<th>Practical Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoke and crave profile</td>
<td>• Tracking device</td>
</tr>
<tr>
<td></td>
<td>• Prompts <em>in-time</em> strategies</td>
</tr>
<tr>
<td></td>
<td>• 5D strategies (delay, distract, drink water, deep breathing, discuss)</td>
</tr>
<tr>
<td>Kick.it stack</td>
<td>• Educational videos and text</td>
</tr>
<tr>
<td></td>
<td>• Distraction stack with 5D strategies</td>
</tr>
<tr>
<td></td>
<td>• Reminders, prompts, and alerts</td>
</tr>
<tr>
<td>Social network platform</td>
<td>• Gamification</td>
</tr>
<tr>
<td>Missions and treatment goals</td>
<td>• Chatrooms</td>
</tr>
<tr>
<td></td>
<td>• Community news and feed posts</td>
</tr>
<tr>
<td></td>
<td>• Links to family, friends, and experts</td>
</tr>
<tr>
<td></td>
<td>• Daily missions</td>
</tr>
<tr>
<td></td>
<td>• Rewards and incentives</td>
</tr>
<tr>
<td></td>
<td>• Logging of pharmacotherapy use</td>
</tr>
</tbody>
</table>

**Study Aims**

This study aimed to gain a deeper understanding of the lived experiences of smoking and quitting among people with SMI and their perception of social support for smoking cessation in relation to the Kick.it app. These experiences informed the researchers understanding of their perceptions regarding the feasibility, utility, and acceptability of the Kick.it app features to guide the tailoring of the app to their specific needs and preferences. The following research questions were investigated: (1) what are the facilitators and barriers to smoking cessation experienced by people with SMI and their perceptions of social support for smoking cessation? and (2) what features of the generic Kick.it app are perceived as feasible, useful, and acceptable in supporting people with SMI to prevent smoking relapse and quit smoking? [14].

**Methods**

**Design**

Co-design methodology was used for this qualitative inquiry because of its value in offering consumer involvement and collaboration [34] in the tailoring of the Kick.it app, which is well matched to meet the study aims [14]. Co-design methods included the triangulation of semistructured in-depth interviews, observation, and in-situ exploration [35] of the Kick.it prototype app with participants.

**Sample**

Research ethics approval was obtained from the Southern Adelaide Clinical Human Research Ethics Committee (reference no. 16.17). A sample of people with SMI was then drawn from the community mental health services (CMHS) and the South...
Australian Cancer Council’s Quitline, within the Adelaide metropolitan area, Australia, between February 2018 and June 2018. Quitline is a free telephone service that provides information and advice to support people to quit smoking [36]. A purposive sampling method was used to recruit participants who were reflective of the target population and who meet the selection criteria [37]. CMHS and Quitline staff identified and screened potential participants’ eligibility to be involved in the study using the selection criteria. The inclusion criteria were (1) self-reported diagnosis of an SMI, (2) adult smokers (aged 18 years or more) who had attempted to quit smoking in the past 12 months and ex-smokers (abstinence for 7 days or longer before the interview) [38], and (3) the ability to provide informed consent as confirmed by CMHS or their doctor. The exclusion criteria were (1) individuals with acute severe suicidality or current acute psychosis as confirmed by CMHS or their doctor, (2) a sensory or motor impairment affecting the individual’s ability to participate in the study, and (3) a severe cognitive impairment affecting the individual’s ability to provide informed consent as confirmed by CMHS or their doctor [15].

Self-reported smoking status was determined using the following smoking status question “Which of the following best describes your smoking status?” and prompted from responses, “I’m a smoker, I smoke daily” and “I’m an ex-smoker, I never smoke now” [39].

CMHS and Quitline staff invited eligible participants to participate in the study and provided the contact details to the research team. Team member (PK) followed up with participants to further describe the study and organize an interview time. At the time of interview, written informed consent was obtained.

**Data Collection**

Two consecutive semistructured in-depth interviews were conducted. This study included an iterative 2-staged interview approach to provide participants with an individualized experience [34]. This was particularly important during the stage 2 interviews, as it enabled people with SMI to receive the personalized assistance needed to navigate the app and the time to reflect and provide feedback on its features.

Stage 1 interviews (approximately 1 hour) consisted of open-ended questions to elicit rich data and a depth of understanding regarding participants’ smoking-related experiences. The interview guide was informed by the research questions, the relevant literature [40,41], and in consultation with the research team. Stage 2 interviews (approximately 1.5 hours) continued to explore participants’ smoking-related experiences in relation to the app. These sessions involved sitting with each participant, as they viewed the prototype app and asking questions in relation to their perception of its features in accordance with the stage 2 interview guide. Observations and field notes were also used to record any reflections on the interview process [37] and the participants’ ability to navigate the app [14]. The stage 2 interview guide was developed from review of the limited studies on tailoring smoking cessation apps for SMI populations [21,24,25], in consultation with the research team, and from preliminary analysis of the stage 1 interviews. The interview guides were reshapen somewhat in accordance with the iterative process of analyzing the data as it was being collected to enable a flexible approach that allows a review and refinement of the interview guide questions (eg, redundant questions were excluded) [37]. Table 1 gives examples of stage 1 and stage 2 interview guides. The interviews were audio recorded and transcribed verbatim by an accredited transcriber to preserve the meaning and authenticity of the participants’ responses. The transcribed interviews were then compared against the audio recordings to ensure their accuracy.

**Data Analysis**

Thematic analysis was used as it provides a systematic approach to organizing, categorizing, and interpreting qualitative data [42,43]. The research team conducted open coding of the first 4 transcripts, independently of each other. In vivo coding was used to exemplify the meaning associated with participants’ responses [44]. A series of team meetings were then held to discuss and debate the initial codes and agree on a structure to guide the coding of the remaining interviews. Categories, selective codes, and emerging themes were captured in a spreadsheet and grouped to assist the researchers to gain a clearer sense of the themes emerging from the data [42-44]. Tables 2 and 3 give examples of stage 1 and stage 2 categories, selective codes, and participants’ frequency of responses.

An iterative process of reading and rereading the transcripts enabled the 3 researchers to reflect on and gain an in-depth understanding of participants’ stories [42,43]. Data were interpreted using a constant comparative approach within and between transcripts to help identify, review, and refine the codes and themes [45]. Mind maps were developed by ordering and linking the codes and categories to the themes. This process prompted robust debate among the researchers, which deepened the interpretation of meaning within the data and finalized the ordering of the themes [42,43]. A dualistic approach was used to utilize existing themes within the literature to build on the limited theory underpinning the design of smoking cessation apps for SMI populations (theory driven) and explore new emerging themes from this study’s findings (data driven) [37,46]. The triangulation of the different research team members’ interpretations and perspectives of the data added further methodological rigor [47]. Sample size was established according to evidence of data saturation being achieved [48]. The researchers concurred that data saturation was achieved by the sixth participant of the stage 2 interviews.
Table 1. Examples of stage 1 and stage 2 interview guides.

<table>
<thead>
<tr>
<th>Interview guide examples</th>
<th>Sample questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Smoking behavior         | • How many years have you smoked cigarettes?  
                          | • What role does smoking play in your life?  |
| Smoking and mental health| • How do you think smoking affects your mental health?  
                          | • What changes do you notice about your smoking when you are feeling psychologically unwell?  |
| Motivation to quit       | • How motivated are you to quit?  
                          | • What motivated you to quit smoking in the past?  |
| Quit smoking attempts    | • During your most recent attempt, what was it like for you to quit?  
                          | • How long did you quit for?  |
| Use of nicotine replacement therapy | • Have you ever used nicotine replacement therapy to assist you to quit?  
                          | • What type of nicotine replacement therapy have you used?  |
| **Stage 2**              |                  |
| App features             | • What do you think about the feature?  
                          | • What do you like/dislike about the feature?  |
| App content              | • How comfortable would you be sharing personal information with the app if it were to lead to a personalized quit program?  
                          | • What do you think about having content specific to mental illness and smoking?  |
| App functionality        | • Can you work out what to do to get to the next screen?  
                          | • What changes are needed to assist people with serious mental illness to work the app?  |
| App aesthetics           | • What do you think about the colors used in the app?  
                          | • What do you think about the font size?  
                          | • What do you think about the quality of the graphic images?  |
| Social support           | • What do you think about talking to other people on the app?  
                          | • What do you like/dislike about social media?  |

*a*Adapted from Vilardaga et al [21], Rotondi et al [24], Vilardaga et al [25], Rand Corporation [40], and Rae et al [41].
## Table 2. Examples of stage 1 categories, codes, and participants’ frequency of responses (N=12).

<table>
<thead>
<tr>
<th>Selective codes</th>
<th>Statistics, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smoking behavior and experiences</strong></td>
<td></td>
</tr>
<tr>
<td>Smoking to manage mental illness/symptoms</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Nicotine addiction</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Stigma associated with smoking/mental illness</td>
<td>5 (42)</td>
</tr>
<tr>
<td>Increased smoking consumption when unwell</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Self-esteem/self-efficacy</td>
<td>10 (83)</td>
</tr>
<tr>
<td><strong>Effects of smoking on mental health</strong></td>
<td></td>
</tr>
<tr>
<td>Perceived benefits of smoking</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Aware of adverse effects of smoking</td>
<td>5 (42)</td>
</tr>
<tr>
<td><strong>Triggers for smoking</strong></td>
<td></td>
</tr>
<tr>
<td>Withdrawals/cravings</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Smoking and mental health</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Smoking to manage life events/stressors</td>
<td>8 (67)</td>
</tr>
<tr>
<td><strong>Quitting behavior and experiences</strong></td>
<td></td>
</tr>
<tr>
<td>Mental illness and smoking relapse</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Difficulty managing withdrawals</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Coping with cravings</td>
<td>11 (92)</td>
</tr>
<tr>
<td><strong>Nicotine replacement therapy</strong></td>
<td></td>
</tr>
<tr>
<td>Use of nicotine replacement therapy</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Never used nicotine replacement therapy</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Positives associated with use</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Adverse side effects of use</td>
<td>6 (50)</td>
</tr>
<tr>
<td><strong>Perceived benefits of quitting</strong></td>
<td></td>
</tr>
<tr>
<td>Saving money</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Improved health</td>
<td>12 (100)</td>
</tr>
<tr>
<td><strong>Barriers to quitting</strong></td>
<td></td>
</tr>
<tr>
<td>Mental illness</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Coping with cravings</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Stress-related factors</td>
<td>10 (83)</td>
</tr>
<tr>
<td><strong>Perception of social support</strong></td>
<td></td>
</tr>
<tr>
<td>Use of social supports</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Reluctance to access</td>
<td>2 (17)</td>
</tr>
<tr>
<td><strong>Use of app/Web-based resources</strong></td>
<td></td>
</tr>
<tr>
<td>Smoking cessation apps</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Health apps</td>
<td>5 (42)</td>
</tr>
<tr>
<td>Other apps (eg, weather)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Never used apps</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Social media (eg, Facebook)</td>
<td>8 (67)</td>
</tr>
</tbody>
</table>
Table 3. Examples of stage 2 categories, codes, and frequency of participants' responses (N=12).

<table>
<thead>
<tr>
<th>Selective codes</th>
<th>Statistics, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>App tailored to app users’ needs</strong></td>
<td></td>
</tr>
<tr>
<td>Creates a profile based on mental illness and smoking</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Develops a personalized quit smoking program</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Tailored strategies specific to mental illness/addiction</td>
<td>10 (83)</td>
</tr>
<tr>
<td><strong>Smoking relapse</strong></td>
<td></td>
</tr>
<tr>
<td>App reassures that quitting can take numerous attempts</td>
<td>6 (50)</td>
</tr>
<tr>
<td>App encourages rapid return to quitting</td>
<td>5 (42)</td>
</tr>
<tr>
<td><strong>An empathetic app</strong></td>
<td></td>
</tr>
<tr>
<td>Uses empathetic/positive communication that looks after self-esteem</td>
<td>6 (50)</td>
</tr>
<tr>
<td><strong>Social network</strong></td>
<td></td>
</tr>
<tr>
<td>Enhance social/peer support for smoking cessation</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Perceived utility of a social network for smoking cessation</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Acceptability of a social network for smoking cessation</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Contingency plan to manage risks/privacy</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Contains chatrooms specific to mental illness and smoking</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Reduces stigma, social isolation, and loneliness</td>
<td>7 (58)</td>
</tr>
<tr>
<td><strong>Kick.it app features</strong></td>
<td></td>
</tr>
<tr>
<td>Utility, usefulness, and acceptability of app features</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Most useful features</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Least useful features</td>
<td>9 (75)</td>
</tr>
<tr>
<td><strong>App functionality</strong></td>
<td></td>
</tr>
<tr>
<td>Able to navigate the app without assistance</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Difficulty navigating the app without assistance (ie, observed usability issues</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Confirmed lack of experience using smartphones/apps)</td>
<td></td>
</tr>
<tr>
<td><strong>App aesthetics</strong></td>
<td></td>
</tr>
<tr>
<td>Colors, font size, and quality of the graphic images</td>
<td>8 (67)</td>
</tr>
</tbody>
</table>

Results

Overview

A total of 12 adults with SMI participated in the study, comprising 6 male and 1 female smokers, and 2 male and 3 female ex-smokers. All participants had been medically diagnosed with either an individual diagnosis of schizophrenia, borderline personality disorder or bipolar disorder, or psychiatric comorbidity. Most participants (75%, 9/12) were diagnosed with paranoid schizophrenia and psychiatric comorbidities, such as depression and anxiety. Some participants (58%, 7/12) had a socioeconomic disadvantaged status, as indicative of these participants’ receiving disability support pension as their primary source of income. All participants were in receipt of community-based support services. Participants’ characteristics are presented in Table 4.

Key findings highlighted several psychosocial factors as important in tailoring the Kick.it app for SMI populations. The key themes that emerged from the data in relation to participants’ lived experiences of smoking and quitting and their perceptions of the Kick.it app features are described below. Key findings aligned with broader psychosocial needs and experiences of perceived stigma and social isolation for this population, which indicated that smoking cessation efforts are inseparable from the environmental and personal context in which these smokers experience and cope with mental illness in their community. Examples of participants’ quotes that help to exemplify the meaning and interpretation of participants’ responses are also included. Multimedia Appendix 5 gives more examples of participants’ quotes.
Table 4. Characteristics of participants with serious mental illness (N=12).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Statistics, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), range (median)</td>
<td>31-53 (47.5)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Female</td>
<td>4 (33)</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>5 (42)</td>
</tr>
<tr>
<td><strong>Smoking behavior</strong></td>
<td></td>
</tr>
<tr>
<td>Heavy smoker (&gt;20, daily)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Years smoked, mean (SD)</td>
<td>26 (12.3)</td>
</tr>
<tr>
<td>Cigarettes smoked per day, mean (SD)</td>
<td>28 (9.9)</td>
</tr>
<tr>
<td><strong>Primary psychiatric diagnosis</strong></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia disorder</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Borderline personality disorder</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Psychiatric comorbidities (n=7)</strong></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>7 (100)</td>
</tr>
<tr>
<td>Depression</td>
<td>4 (57)</td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>1 (14)</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>1 (14)</td>
</tr>
<tr>
<td><strong>Level of education</strong></td>
<td></td>
</tr>
<tr>
<td>Tertiary education</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Technical and Further Education</td>
<td>4 (33)</td>
</tr>
<tr>
<td>High school</td>
<td>6 (50)</td>
</tr>
<tr>
<td><strong>Source of income</strong></td>
<td></td>
</tr>
<tr>
<td>Full-time work</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Part-time work</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Disability support pension</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Partnered</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Divorced</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Type of residence</strong></td>
<td></td>
</tr>
<tr>
<td>Supported residential facility</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Independent living</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Living alone</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Living with others</td>
<td>6 (50)</td>
</tr>
<tr>
<td><strong>Quit attempts</strong></td>
<td></td>
</tr>
<tr>
<td>Single attempt</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Multiple attempts</td>
<td>9 (75)</td>
</tr>
</tbody>
</table>
### Use of social support resources

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Statistics, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family and friends</td>
<td>8 (67)</td>
</tr>
<tr>
<td>General practitioner</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Mental health caseworker</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Quitline call center</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Nicotine replacement therapy</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Smoking cessation apps</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Smartphone ownership</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Use of social media</td>
<td>8 (67)</td>
</tr>
</tbody>
</table>

*A total of 7 participants presented with psychiatric comorbidities.*

### Special Needs

**An App That Tailors a Personalized Quit Program to an Individual's Psychosocial Needs**

Exploration of the participants’ perception of the Kick.it app features highlighted the importance of the app tailoring a personalized quit program to their needs. This included the app tailoring a program specific to their psychiatric diagnoses (and consequent symptoms) and smoking behavior. Participants described their smoking behavior as a form of self-medication as it provides them with a source of comfort to relieve their symptoms of mental illness:

...an app that’s tailored to mental health consumers is essential...if it’s generic it won’t delve into the personal struggles that they’re going through with having to look at smokes as being their only source of comfort. [interview session (IS) 1, participant (P) 1]

All participants indicated that their smoking consumption almost doubled when they were feeling psychologically unwell, which reflected their reliance on cigarettes to help them to cope with their mental illness:

When I’m depressed...I just would like to be left alone with my cigarettes and coffee...[smoking] goes up to about 40 a day. [IS1, P3]

Some participants (83%, 10/12) also indicated that stressful social environments, such as relationship problems, peer smoking, and work-related issues, were barriers to smoking cessation:

I use it as a stress reliever...where I’m completely thinking of nothing else other than smoking...I’m not worried about uni, family or work problems... [IS1, P7]

It was evident from listening to participants stories that people with SMI have many psychosocial issues and need support to manage their mental illness, nicotine dependency, and social-related issues while attempting to quit (see Multimedia Appendix 5 for more quotes relating to this theme).

### An App That Normalizes Smoking Relapse and Multiple Quit Attempts

Exploration of participants’ quitting experiences revealed that most participants had attempted to quit on several occasions (75%, 9/12), but their attempts were often short lived. For example, some participants’ recalled occasions where they were determined to quit, but within a few hours when the intensity of the cravings had occurred, they were reaching for a cigarette:

I have [tried to quit] many times. Two hours later I’ve got a fag in me hand. [IS1, P3]

Participants described overwhelming feelings of disappointment and helplessness regarding their ability to sustain a quit attempt. These findings indicated that an app tailored to support people with SMI may focus more on reassuring them that smoking relapse is a normal part of the quitting process, and that it can take numerous attempts to quit:

When you relapse you’re disappointed with yourself and you smoke more than you did before...the best thing about an app that reassures you is that...it’s okay to have the relapse but get back on the bandwagon...try the app again. [IS2, P1]

Our findings also indicated that standard smoking cessation approaches that require a range of cognitions such as critical and analytical thinking, evaluating, judging, and weighing options, and deciding on actions that can foreground planning to quit caused participants heightened anxiety and stress. Participants reported that feelings of anxiety and stress increased their smoking consumption and were major barriers to smoking cessation. Therefore, asking people with SMI to recall quit strategies in those moments when they are feeling anxious and experiencing intense withdrawals offer limited smoking cessation support. These findings indicated that there is a need for alternative smoking cessation approaches that address these temporal issues by assisting people with SMI to quit smoking in real time, within the context of their daily lives:

If you can offer practical solutions for people to try in certain situations...that would be a much better deterrent to lighting up. [IS2, P7]

Participants were impressed with the smoke and crave profile generated in the Kick.it app. They perceived the *in-time* quit...
strategy messages that app users receive when they log a smoke or crave as useful and acceptable in supporting them to quit. They also liked the tracking device as it would provide them with ongoing feedback regarding their smoking and quitting behaviors (see Multimedia Appendix 5):

*Might be doing something then all of a sudden you get a message and you think...I’ll give that a try.* [IS2, P9]

**Strong Focus on User Experience to Improve Usability of the App**

Findings suggest the need to apply an optimal UX design through simple user interfaces [27] such as directional cues with arrows indicating to *swipe here* to improve usability of the app among people with SMI:

*I didn’t know what to do when I was sliding across...you’re going to need a sign to say slide across here. It has to be really basic for people who are mentally ill.* [IS2, P6]

Furthermore, 2 participants with schizophrenia found navigating the app overwhelming as they possessed limited knowledge and skills in technology. Therefore, applying a simple app design increases the likelihood that people with SMI will be able to use the adapted Kick.it app:

*What is it, an app? What does that mean? A phone? Email, it’s got internet on it? I’m not really quite sure what’s going on.* [IS2, P2]

Participants also reported that they appreciated our co-design approach as it enabled them to share their smoking-related stories and provide input on the tailoring of the app for SMI populations (see Multimedia Appendix 5):

*The fact that you are interviewing me and other people with different experiences...you are making it [the app] really consumer-focused.* [IS2, P12]

**Preferences**

**A Caring App**

Participants wanted a caring app, with almost human-like qualities that could offer companionship and enable them to share their concerns without feeling stigmatized or judged:

...if they get the idea that people actually care about the cigarette smoking...that people actually care for their health. [IS2, P3]

The need for a caring app seemed to stem from the interplay between SMI, smoking, and stigma that featured heavily in both interviews and were common experiences among all participants. For example, some participants talked about schizophrenia being less accepted in the community than depression, which resulted in them experiencing social isolation associated with not having friends. Some participants also indicated that stigma was a major force driving their smoking behavior:

...people with schizophrenia get pushed away, and that’s why they get into their circles...smoking cigarettes...there’s a lot of stigma especially with things like schizophrenia. [IS2, P3]

In addition, participants perceived that a caring app could contain messages that motivated them to quit, to believe in themselves, and that gave them hope that they could quit smoking (see Multimedia Appendix 5):

*It’s nice to receive a positive statement because it’s quite daunting quitting, and you feel quite alone and isolated, like can I do this.* [IS2, P7]

**A Social Network–Based App**

Most participants (83%, 10/12) were enthusiastic about engaging with a social network. They liked the idea of having chatrooms specific to mental illness and smoking where they could connect with likeminded people who also wanted to quit. By enhancing peer support for smoking cessation through the social network function, the app has the potential to address stigma, social isolation, and loneliness. People with SMI can use their phone anytime and anywhere to connect with other people who are also using the app:

*What’s good about it is following each other and giving each other support...they can interact with each other, because it’s important.* [IS2, P5]

In relation to privacy and confidentiality, most participants liked the inclusion of *terms and conditions* that outlined the privacy settings and *rules of use* to alleviate potential concerns around engaging with the social network (see Multimedia Appendix 5):

...it’s bound by privacy so you know you can talk about this issue...and it’s just the community that you’re working on this issue with...it’s not going out to everybody. [IS2, P12]

**Social Support Resources for Smoking Cessation**

Exploration of participants’ perception and utility of social support for smoking cessation indicated that most participants received social support (83%, 10/12) from their family, friends, and general practitioner:

*I’d planned to give up smoking with a friend...mum was supportive...* [IS1, P4]

Many participants (75%, 9/12) reported using nicotine replacement therapy to assist their quit attempts, but some participants had experienced adverse physical and/or mental health effects (50%, 6/12), which deterred them from continuing its use:

...you still feel like smoking on nicotine replacement therapy, but the cravings are not as bad, you don’t get as agitated without smokes. [IS1, P8]

Overall, 3 participants reported that they had used a smoking cessation app to support a quit attempt. Of these, 2 participants found the app useful in supporting their attempt. It is often assumed that younger people are more likely to use apps; however, this finding provided some insight into the use of smoking cessation apps among adults aged 36 to 52 years, with SMI (see Multimedia Appendix 5):
It’s [the app] just something that’s always there...it was there for me at the touch of a phone. [IS1, P4]

The key findings guiding the tailoring of the Kick.it app for SMI populations, including participants’ lived experiences of smoking and quitting, their identified needs and preferences for tailoring the Kick.it app, and the smoking cessation pathway outlining the smoking relapse cycle and smoking cessation intervention, are illustrated in Figure 2.

**Figure 2.** Key findings guiding the tailoring of the Kick.it app for serious mental illness populations.

### Mental illness and smoking

#### Inextricable link
- Coping
- Identity
- Belief systems
- Sense of worth

#### Psychosocial factors
- Being a person with mental illness
- Complex issues
- Nicotine addiction
- Cognitive deficits
- Social inclusion

#### Barriers
- Mental illness
- Nicotine dependency
- Social exclusion
- Stress
- Habit
- Boredom

### App adaptation functions

#### Special needs
- Personalized quit program
- Normalization of smoking relapse
- Strong focus on usability

#### Preferences
- Caring app
- Social network

### Smoking cessation pathway

#### Smoking relapse cycle
- Multiple quit attempts
- Ability to sustain a quit attempt
- Solutions to facilitate sustained smoking abstinence

#### Interventions
- In-time strategies
- 5D strategies
- Pharmacotherapy education and goal setting
- Social support education and strategies eg, social network chatrooms and community feeds
- Tracking device
- Gamification for quit skills development

### Discussion

#### Principal Findings and Comparison With Previous Work

Results suggest that tailoring of the Kick.it app is feasible, useful, and acceptable for people with SMI. Participants were interested in using mHealth interventions to support their quit smoking efforts. Similar to other studies [21,25], our findings indicated using a co-design approach to improve the utility, usability, and acceptability of smoking cessation apps for SMI populations is required.

Exploring participants’ smoking and quitting experiences highlighted the complexity associated with the inextricable link between mental illness and smoking [6,7], which has pervasive impacts on participants’ psychosocial well-being and affects every aspect of their daily lives. These findings suggest that specific tailoring of the app is needed to assist people with SMI to navigate the complex interaction between mental illness and smoking that impact on their capacity to quit. Consistent with previous evidence [10], results of this study found that current smoking cessation approaches are limited in supporting people with SMI to quit as they do not account for their mental health–related needs [49], nor do they address the symptoms of mental illness and nicotine dependency simultaneously. The need for a dual approach to smoking cessation that provides quit strategies to address both the symptoms associated with smokers’ psychiatric diagnoses and nicotine addiction was identified many years ago [50]. Another limitation of current smoking cessation approaches is that they do not offer real-time assistance regarding experiencing the urge to smoke. For example, they do not address stressors arising in the context of their daily lives in the moments [33] when those stressors are heightened, and the person is at high risk of reaching for a cigarette to alleviate their distress. The delivery of novel smoking cessation approaches that offer assistance in the here and now to support people with SMI to prevent smoking relapse and quit smoking has been established as important in this study and warrants further investigation. The Kick.it app’s smoking cessation interventions may provide the solution to address the limitations of current smoking cessation approaches and support people with SMI to quit.
This study also provided valuable insight into the effects of stigma and a possible solution to the perpetuated entrenched marginalization and social disadvantage among smokers with SMI [11,49]. Gaining a deeper understanding of participants’ lived experiences of stigma and how those experiences had impacted on their self-worth highlighted why it was important to them to have a caring app. Therefore, the features of the Kick.it app may be different from generic smoking cessation apps in that it could focus more on building their self-esteem and self-efficacy. There is also potential for the Kick.it app’s caring features and social network to reduce the effects of stigma, social isolation, and loneliness by enhancing social inclusion, and a sense of belonging to a social support network for smoking cessation. The benefits of using a social network to gain peer support for smoking cessation among people with psychosis has been established [51].

An adapted Kick.it app has the potential to become an all-encompassing solution, a virtual friend that offers around the clock support to help people with SMI to quit, and address stigma which is a by-product of SMI [49]. The specific app features and their adaption that participants perceived would be feasible, useful, and acceptable in assisting SMI populations to quit smoking included the smoke and crave profile, the supportive messaging, the tracking device, and social networking.

Limitations
Limitations of the study included the participants’ self-reported smoking status and psychiatric diagnoses. However, smoking status was obtained using a screening tool adapted for this population [39], and all but 1 participant was recruited from CMHS, which provides ongoing case management support to people with existing SMI. Another limitation relates to the participants’ sampling a prototype of the Kick.it app at the stage 2 interviews which provided users with a brief window of 1.5 hours to view the app and advise on its features in an interview setting rather than a naturalistic setting [37].

Conclusions
This study provides evidence for innovative smoking cessation approaches to support people with SMI to prevent smoking relapse and successfully quit. We contributed to the limited knowledge on designing smoking cessation apps for SMI populations by using a co-design principal based on the IM framework to explore their lived experiences of smoking and quitting and their perception of the Kick.it app features to guide the tailoring of the app. Through the lens of people with SMI, this study provides insight into the smoking behaviors and personal struggles they encounter in their endeavor to quit smoking. By confronting some of the major barriers to smoking cessation for this population, this study contributes to possible solutions for important mental health–related issues, including stigmatization and social isolation [11,49]. The next stage of research planned by the authors of this study involves tailoring the Kick.it app in accordance with the findings and then conducting a quantitative study to gain a representative sample to assess the effectiveness, utility, and acceptability of the app among people with SMI in relation to smoking cessation [14].

Acknowledgments
This research is supported by a research seeding grant from Flinders University. Grant funds do not cover the cost of open access publishing of this study. This study makes an original contribution that forms part of a larger study design protocol for adapting the Kick.it app for SMI populations, published by Lawn et al. The authors would like to thank the people who participated in the study for their contributions on tailoring of the Kick.it app for SMI populations, staff of CMHS, and the South Australian Quitline who supported participant recruitment, James Stewart, the wider Kick.it team, and the Hospital Research Foundation for their role in establishing the Kick.it app.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Screenshot of a social support in-time intervention.
[ PNG File 282 KB - humanfactors_v6i3e14023_app1.png ]

Multimedia Appendix 2
Screenshot of a 5D strategy in-time intervention.
[ PNG File 370 KB - humanfactors_v6i3e14023_app2.png ]

Multimedia Appendix 3
Screenshot of the tracking device.
[ PNG File 137 KB - humanfactors_v6i3e14023_app3.png ]

Multimedia Appendix 4
Screenshot of an educational video on nicotine replacement therapy.
References


Abbreviations

CMHS: community mental health services
IM: intervention mapping
mHealth: mobile health
SMI: serious mental illness
UX: user experience

Edited by A Kushniruk; submitted 14.03.19; peer-reviewed by C Mather, S Robins, E Borycki; comments to author 20.05.19; revised version received 11.07.19; accepted 29.07.19; published 03.09.19.

Please cite as:
Klein P, Lawn S, Tsourtos G, van Agteren J
Tailoring of a Smartphone Smoking Cessation App (Kick.it) for Serious Mental Illness Populations: Qualitative Study
JMIR Hum Factors 2019;6(3):e14023
URL: https://humanfactors.jmir.org/2019/3/e14023
doi: 10.2196/14023
PMID:31482850

©Pauline Klein, Sharon Lawn, George Tsourtos, Joep van Agteren. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 23.09.2019 This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.