Contents

Original Papers

Development and Implementation of the Portable Operating Room Tracker App With Vital Signs Streaming Infrastructure: Operational Feasibility Study (e13559)
Matthias Görges, Nicholas West, Christian Petersen, J Ansermino. ................................................................. 2

A Postoperative Pain Management Mobile App (Panda) for Children at Home After Discharge: Usability and Feasibility (e12305)
Dustin Dunsmuir, Helen Wu, Terri Sun, Nicholas West, Gillian Lauder, Matthias Görges, J Ansermino. .......................... 16

Comparing the Relaxing Effects of Different Virtual Reality Environments in the Intensive Care Unit: Observational Study (e15579)
Stephan Gerber, Marie-Madlen Jeitziner, Simon Sänger, Samuel Knobel, Laura Marchal-Crespo, René Müri, Joerg Schefold, Stephan Jakob, Tobias Nef. .......................................................................................... 27

Impact of an Intensive Care Information System on the Length of Stay of Surgical Intensive Care Unit Patients: Observational Study (e14501)
Camille Havel, Jean Selim, Emmanuel Besnier, Philippe Gouin, Benoit Veber, Thomas Clavier. ........................................ 36

Listening to the HysterSisters: A Retrospective Keyword Frequency Analysis of Conversations About Hysterectomy Recovery (e10728)
Arpit Dave, Johnny Yi, Andy Boothe, Helene Brashear, Jeffrey Byrne, Yash Gad. .......................................................... 44
Development and Implementation of the Portable Operating Room Tracker App With Vital Signs Streaming Infrastructure: Operational Feasibility Study

Matthias Görges1,2, PhD; Nicholas C West1, MSc; Christian L Petersen1,2, PhD; J Mark Ansermino1,2, MBBCh, MMed (Anaesth), MSc (Info), FRCPC

1Department of Anesthesiology, Pharmacology & Therapeutics, The University of British Columbia, Vancouver, BC, Canada
2Research Institute, BC Children’s Hospital, Vancouver, BC, Canada
3ESS Technology Inc, Kelowna, BC, Canada

Corresponding Author:
Matthias Görges, PhD
Department of Anesthesiology, Pharmacology & Therapeutics
The University of British Columbia
950 W 28th Avenue, Rm V3-324
Vancouver, BC, V5Z 4H4
Canada
Phone: 1 604 875 2000 ext 5616
Fax: 1 604 875 2668
Email: mgorges@bcchr.ca

Abstract

Background: In the perioperative environment, a multidisciplinary clinical team continually observes and evaluates patient information. However, data availability may be restricted to certain locations, cognitive workload may be high, and team communication may be constrained by availability and priorities. We developed the remote Portable Operating Room Tracker app (the telePORT app) to improve information exchange and communication between anesthesia team members. The telePORT app combines a real-time feed of waveforms and vital signs from the operating rooms with messaging, help request, and reminder features.

Objective: The aim of this paper is to describe the development of the app and the back-end infrastructure required to extract monitoring data, facilitate data exchange and ensure privacy and safety, which includes results from clinical feasibility testing.

Methods: telePORT’s client user interface was developed using user-centered design principles and workflow observations. The server architecture involves network-based data extraction and data processing. Baseline user workload was assessed using step counters and communication logs. Clinical feasibility testing analyzed device usage over 11 months.

Results: telePORT was more commonly used for help requests (approximately 4.5/day) than messaging between team members (approximately 1/day). Passive operating room monitoring was frequently utilized (34% of screen visits). Intermittent loss of wireless connectivity was a major barrier to adoption (decline of 0.3%/day).

Conclusions: The underlying server infrastructure was repurposed for real-time streaming of vital signs and their collection for research and quality improvement. Day-to-day activities of the anesthesia team can be supported by a mobile app that integrates real-time data from all operating rooms.


KEYWORDS
communication systems; patient monitoring; user-centered design; human factors; anesthesia
Introduction

Background

Pediatric perioperative care can be a hectic and stressful work setting, in which a multidisciplinary team of clinicians (anesthesiologists, surgeons, nurses, and other health professionals) continually observes and evaluates patient information [1-3]. Efficiency and patient safety in the procedural suites depend on a well-functioning team [1,2]; however, data are often limited to certain locations, cognitive workload can be high, and team communication constrained by availability and differing priorities [4,5]. Even experienced teams face challenges of physical separation and the need to locate each other by pager and phone [6].

To address this problem, our team of clinicians, engineers, and computer scientists developed a remote Portable Operating Room Tracker app (the telePORT app) to improve information exchange and communication between anesthesia team members. It aims to combine a real-time feed of physiological waveforms and vital signs from the operating rooms (ORs) with messaging, paging and reminder functions.

Clinical Setting

BC Children’s Hospital (BCCH) is a tertiary pediatric medical center in which approximately 11,500 children per year undergo general anesthesia for surgery, dental procedures, endoscopic investigations, medical imaging, and other interventions. On any weekday, there are approximately 12-15 anesthesiologists covering a variety of locations, including the core 8 ORs, the oncology suite, X-ray, computed tomography (CT) scan, magnetic resonance imaging (MRI) and ultrasound rooms, and a burn treatment room.

Their work is supported by a small team of anesthesia assistants (AAs). These are allied health professionals, with backgrounds in respiratory therapy, who work with the attending anesthesiologists to optimize the safety of the anesthetized children by ensuring that equipment is maintained and readily available, and by providing extra assistance when needed. They are a limited resource, as only 3-4 AAs are available at any time and they are often tied up helping start complex procedures, such as cardiac, neuro-, or spine surgery. Each AA typically supports more than one anesthetic location and ideally needs to be able to monitor the status of multiple patients. They must be responsive to requests for assistance and, if possible, maintain ready awareness of the situation that precipitated the call for help.

To request an AA, anesthesiologists relied on a numeric, one-way, phone-based paging system, which could be cumbersome to use from the anesthesia workstation, provided no feedback about AA availability or the urgency of the message, did not allow two-way communication, and did not allow delegation of tasks if the AA was unavailable. Nonetheless, the paging approach was well established, as it was secure and reliable and was still in use at BCCH when we began the telePORT project. It gave us the opportunity to improve teamwork through two-way, secure communication and information to improve the situational awareness of the AAs.

Technical Setting

Physiological multiparameter patient monitors in ORs and intensive care units (ICUs) gather data from multiple sensors and diverse formats. These include, for example, physiological waveforms like the electrocardiogram (ECG), from which numeric data such as heart rate and alarms (eg, bradycardia [low heart rate]) are derived. Most ORs and ICUs are multibed environments, in which individual patient monitors are connected to central monitoring stations that support event recording, centralized printing and remote alarm monitoring. These central stations are immobile and have limited proprietary interfaces, which impose a barrier to secondary use of their data. Yet significant opportunities exist if these data can be made available on a secure mobile platform, ideally augmented with decision support systems.

These data gathered from vital signs, covering a wide range of settings and procedures, are valuable for research, quality improvement (QI), and review of undesired clinical events or outcomes that occur during normal clinical care.

At BCCH, multiple vital sign parameters are monitored in the OR (including the postanesthetic care unit) and pediatric ICU, and they are captured automatically and stored in a central station server. Yet, this mass of data is routinely discarded after 2-3 days, with only small samples manually or automatically transcribed in infrequent intervals (typically at 5 min, 15 min, 1 hr, or 4 hrs) into the patient’s medical record. To overcome this limitation, we obtained ethical approval to collect and store all vital signs data from children in the OR in 2009 and in the ICU in 2011 for research and QI purposes.

LambdaNative is an open source software framework, designed for developing medical apps, that promotes deterministic, robust and correct code [7]. It is based on the portable Gambit Scheme programming language [8] and provides a flexible cross-platform environment for developing graphical apps on mobile devices as well as medical instrumentation interfaces on embedded platforms [7]. It has been used to create a diverse range of applications, from a mobile data collection tool for a multicenter preeclampsia trial in low resource settings [9], to mission critical, embedded drug delivery as part of a closed-loop anesthesia system [10]. Importantly, we identified strategies to improve medical device data exchange using a data broker concept [11].

Aim of Study

In this article, we describe the development of the telePORT mobile app using user-centered design. We also outline the back-end infrastructure (VitalNode) required to extract monitoring data, facilitate data exchange, and ensure privacy and safety. This section includes some technical details that may be helpful to some readers (in particular those who do not have an Anesthesia Information Management System).
Methods

The Main App Features

The purpose of telePORT is to improve information exchange and simplify communication between anesthesia team members. We focused the design on five key features, all of which could be reached at any point using a menu navigation bar at the bottom of the screen (Figure 1).

Overview Screen

This screen provides basic information about the patient monitoring locations to which the user is subscribed. The information displayed includes an anesthetic phase indicator and three vital signs: heart rate (HR), oxygen saturation (SpO2), and end-tidal carbon dioxide concentration (etCO2). A messaging icon displays the number of unread notifications for each location (Figure 1). Secondary screens show either waveforms and additional numeric values, in a design similar to the patient monitor layout, or a 30-minute trend screen, which can be obtained by selecting a location (Figure 1). The monitoring screen allows swiping to switch quickly between screens and features a full screen mode which can be reached by changing the orientation of the device to landscape. This was added late in the development cycle as a frequently requested feature.

Messaging Screen

This screen provides a combination of person-to-person chat and a system to receive requests for help through a button pressed on the patient monitor in an OR. Both types of message are located in two separate areas of the screen with automatically generated messages (requests for help from the OR and notifications from reminders) in the top panel and sent messages in the bottom panel (Figure 1).

Figure 1. Progression of the telePORT app from initial concepts (top row), to intermediate prototype re-designed after user feedback (middle row), to final app (bottom row). The app features shown are: room subscription screen (first), overview screen (second), waveform/numeric detail screen (third), help request screen (fourth), reminder setup (fifth), and messaging overview screen (sixth).
Requests for help from the OR are initiated by the anesthesiologist pressing the physical Snapshot button on the patient monitor, with two presses in a 4 second window initiating an urgent (STAT) call. These requests prompt the display of a popup window. If selected from the popup or system messaging list, further information is displayed including the origin location, message urgency (“Please come soon” or “STAT Request”), a selection of vital signs including HR, SpO$_2$, etCO$_2$, respiratory rate (RR), and blood pressure (NIBP). The app allows a response of Okay or Ignore (Figure 1). If the user presses Ignore or fails to confirm the message within 5 min, the alarm is escalated to all other users logged into the system. Urgent (STAT) messages are sent to all users, regardless of the locations to which they are subscribed. Providing vital signs with the request allows the user receiving the message to obtain some information about the situation before arrival, in an effort to raise situational awareness [1,12,13].

For the person-to-person messaging system, we support quick texts, with 10 options including messages such as “Yes,” “No,” “Can you help?,” “Meet in office,” and so on, as well as a regular two-person chat system with an on-screen keyboard. The quick text list was developed by asking anesthesia assistants for the most commonly used texts rather than through something like a self-learning approach. A special user (ALL) allows messages to be sent to all users, but responses will only reach the sender as group chat was not implemented.

Phonebook Screen
This screen lists frequently used phone extensions, such as each OR, the biomedical engineering department and pager numbers, and can be edited within the app. We envisaged adding a Voice over internet protocol (VoIP) feature to this screen but it was never completely implemented.

Location Subscription Screen
This screen shows which locations a user is subscribed to and also the other team members requesting information from that location. The user can subscribe to, or unsubscribe from, each of the monitoring locations and subscriptions can be delegated to other users (the recipient must accept the transfer). The initial design called for a visual map, with icons for cardiac, dental, neurosurgical, orthopedic and other ORs, to allow users to select locations based on their physical proximity. It was later modified into a simple list as it is faster to use, and as assignments are more frequently based on the case complexity and not physical proximity.

Reminder Screen
The fifth screen allows the user to set up reminders based on time or anesthetic phase (Figure 1). For time-based reminders, the user selects time (using scrolling wheels for hours, and minutes in 5 min increments), reminder frequency (none, 15, 30, 60, 120, and 300 min), and tasks (“Check Room,” “Draw ABG,” and “Meeting”). For reminders based on anesthetic phase, only the location, task, and anesthetic phase symbol are selected. A special empty room cycling mode allows repeated reminders whenever the room phase changed to empty until the end of the day. It was intended for use in high turnover lists, such as endoscopies, where restocking is performed frequently throughout the day. Repeated reminders were a requested feature to support the arterial blood gas (ABG) workflow, in which the anesthesiologist periodically (eg, every 2 hrs) draws a blood sample, which the anesthesia assistant then takes to the point-of-care analysis device for processing before returning with the results to the OR.

Development of the VitalNode Server Component
The multiparameter patient monitors transmit real-time data to a central station via a dedicated ethernet network. To overcome limited access to the proprietary central station interface, we intercepted raw ethernet data packages directly.

We initially used a passive network tap on the transmitting wires of the ethernet cable connecting the main OR and ICU switch aggregating data from the monitor to the central station. These were connected to a network port which was set to monitoring (promiscuous) mode to collect the received network traffic. In a later iteration we bridged two network ports, as this solution was easier to implement on the updated 100-megabit full duplex network which automatically assigned receiving or transmitting wire pairs. The device doing this work was called ‘VitalNode.’ A hybrid approach was taken for the data, with the raw data stored but also parsed in real-time for processing (Figure 2).
Figure 2. Overview of the data collection system. Data from the patient monitors is captured by the VitalNode box before it flows through the switch to the central station. This data is parsed, encrypted, and made available to the telePORT app for real-time use, and also stored in both parsed (trends and waveform CSV files) and raw formats. PCAP: packet capture.

Real-Time Data Parsing
Real-time parsing of the network traffic involved decoding the data frames of the Philips and GE or Datex monitor communication protocols. The Datex protocol was published accurately in the programming guide, whereas the Philips protocol had undocumented header data structures that were decoded using empirical analysis. For performance reasons we chose to employ a kernel level filter, restricting parsing only to packages of types we knew the parser could analyze and which contained data useful for our purposes.

The VitalNode app performed several tasks, including: 1) extracting key physiologic trend data from all available vital signs (e.g., SpO\textsubscript{2}, HR, temperature) at 10 sec (Datex monitors) or 1 sec time resolution (Philips monitors); 2) parsing alarm data; 3) parsing relevant physiological waveform data (e.g., ECG, plethysmograph) at resolutions between 25 and 500 Hertz depending on the vendor and variable; and 4) extracting demographic information (rarely entered).

These data were written to disk in comma-separated values (CSV) format for later use, separating the 5 or 10 second interval trend files of all potential variables that could be captured. Then, waveform files were written if a sensor was sending valid data, to simplify data extraction and processing. Data was kept in memory for near real-time data access in streaming applications. VitalNode also stored the messages exchanged in the telePORT app, both by provider (to allow person-to-person communication), and by monitor location (to store system generated messages that were available to all subscribers of a location).

The app aggregated 30 mins of trend data (in 10 sec resolution), for ease of display. It also applied some simple data processing, including an anesthesia phase indicator. This used a simple 4-state machine, based on the presence or absence of valid values from three commonly used sensors (SpO\textsubscript{2}, HR, and etCO\textsubscript{2}, using 30 sec averages). This allowed us to determine if the case had started (induction of anesthesia, one sensor present), was in the middle of the case (maintenance of anesthesia, all three sensors present), was nearing the end of the case (emergence from anesthesia, one sensor off), or if the room was empty (no sensors reporting valid data). To avoid problems during cardiopulmonary bypass, zero values for etCO\textsubscript{2} were considered valid while empty or error states were deemed invalid.

Raw Data Storage
We used raw network package capture (PCAP) files in 10-min chunks, compressed using bzip2, for two reasons: 1) these files use disk space efficiently; and 2) this approach allowed us to capture data elements not implemented in the data parser and to recover from potential errors in the parser code.

Development of the telePORT Client Component
A user-centered design process was used [14]. Initially, three AAs were shadowed during two of their day shifts to identify their information needs, communication and reminder strategies, and task planning approaches. Combined with semi-structured interviews, these data were used to create a modified work-domain analysis [15,16], which established a hierarchical model of the domain and allowed specification of the app
requirements. Next, we conducted a participatory design of the mobile app [17-19], in which we sought frequent feedback on design features, initially using mockups developed in PowerPoint (Microsoft, Seattle, WA), and later in partially working prototypes presented on an iPod touch (Apple, Cupertino, CA; Figure 1).

We used LambdaNative [7] to rapidly develop a mobile app targeted to run on a small mobile tablet or smart phone, such as a 4th generation iPod touch, with an effective screen resolution of 480x320 pixels. Initial development was performed on a Linux machine with only the iOS building process performed on Mac, thus highlighting one advantage of the cross-platform capabilities of the LambdaNative environment. The prototype components were implemented, underwent quick periods of usability testing by the AAs as well as research team members, and then were modified according to the feedback obtained. The app was deployed on iPod Touches using an Apple developer provisioning profile. A modified version (without continuous background processing and connectivity) is available on the Apple App store [20].

**Evaluation Approach**

With approval by the University of British Columbia Children’s & Women’s Health Centre of British Columbia Research Ethics Board (H11-01785) and written informed consent, we established the baseline workload of the AAs using a hybrid approach: we used a step counter to obtain a surrogate for distance travelled, as well as daily logs of the number of pages received and the urgency of those requests. We originally planned to repeat these after the telePORT app was implemented, but the data was highly variable and data collection was not always comprehensive as users considered it too onerous, so we decided not to.

To evaluate the use of the telePORT app, we used both formal and informal feedback from AAs by email or verbal messages and quantified usage patterns using VitalNode-based data logging. Time stamped data of user logins, logouts, disconnects (loss of client connections exceeding 60 secs), navigated screens, message sender and recipients, and system-generated pages were recorded in CSV format on the VitalNode server. Data were parsed using R (R Foundation for Statistical Computing, Vienna, Austria), and plotted for analysis. We used battery usage as a surrogate for device usage (it being removed from the charging station for use) and explored message frequency and number of pages quantitatively. Cytoscape (Institute of Systems Biology, Seattle, WA) was used in an attempt to map navigational flow by creating bidirectional graphs of app screen log sequences.

**Results**

**Baseline Step and Pager Tracking**

In total, 27 data sheets were completed from October to December 2011 (an estimated response of 10-15% of shifts worked). The median (interquartile range [IQR]) number of steps was 6426 (IQR 5163-8910) per shift, or 627 (IQR 451-784) per hour. During this time the AAs reported 34 pages, which is a median of 0 (IQR 0-2.5) per shift, with the reasons for each including equipment in 12/34 (36%) cases, clinical in 14/34 (41%) cases, and administrative in 8/34 (24%) cases.

**Battery and System Usage**

An early version of telePORT drained the iPod touch (4th generation, iOS 5.1, screen switched off) in 6-7 hours, which was insufficient for clinical use. Code was optimized to lower central processing unit (CPU) usage, achieving average battery drainage of 9.04% per hour, which allowed for 11 hours of battery life between charges.

During the evaluation, telePORT was used exclusively by the AAs (not by the anesthesiologists or Anesthesiologist in Charge). Data recorded on the 337 days between January 7, 2013 and December 9, 2013 were used to analyze usage, as this period coincided with peak usage and was available continuously. Battery drainage (as a surrogate of app use) showed usage patterns that overlapped with expected use, primarily during weekday OR hours, starting at 7h, with peak usage between 9h and 17h, and continuing until about 19h (Figure 3).

While usage varied widely during the evaluation period, with a median (IQR [range]) of 1.45 (1.0-2.0 [0.01-3.7]) devices in use each day, there was some continual decline in usage (~0.3% of usage per day) towards the end of the observation period. This decline was partly attributed to network communication issues reported by users but might also indicate workflow integration issues.
Communication Flow and System Navigation

A total of 327 messages (approximately 1/day) and 1528 requests for help (approximately 4.5/day) from an OR were recorded. Messaging pattern analysis found communication flow centered on one AA (AA5), who was not the team leader (Figure 4). The distribution of requests for help showed that OR7 (spine surgery) outnumbered other locations, but there was also a discernible pattern in OR1 which peaked on Tuesdays and Fridays (typically neurosurgery). Timing distributions did not show any surprising patterns, other than a peak number of messages in the middle of the day when new cases started (Figure 5).

The most commonly used features were Messaging (30%), Overview (20%), Pages or Alerts (19%), and Waveforms (14%). The integrated Chat screen was rarely used (2%), indicating a preference for using the 10 quick texts over free text communication. As the navigation bar allowed instant navigation between the five main features, only subscreen navigation shows signs of directed navigational patterns (eg, going from Reminder setup to Reminders).
Figure 4. User communication pattern within the telePORT app. Blue nodes represent the anesthesia assistants and green nodes are other roles. The width of the connections is proportional to the number of messages sent. Black arrows indicate the receiver of the message. AA: anesthesia assistant.
Figure 5. Paging feature utilization, split by weekday. Left panel shows paging usage by location, with those occurring inside core operating room hours indicated in mint and those outside these hours in light purple; right panel has data by time of day. Overlaid are operating room core hours using black dashed lines and typical user shift hours using red dashed lines.

Discussion

Main Findings

We developed the telePORT app to improve information exchange and communication between anesthesia team members, including messaging, help request (paging), and reminder features. In addition, telePORT provided a real-time feed of waveforms and vital signs from the ORs. Following development of the app’s user interface with user-centered design principles, telePORT was used by AAs at BCCH over 11 months to support their normal day-to-day activities. There were more help requests from anesthesiologists in the ORs (approximately 4.5/day) than messages between team members (approximately 1/day), but passive OR monitoring was also used frequently (34% of screen visits). Intermittent loss of wireless connectivity led to a steady decline in usage (~0.3%/day) and the project was eventually halted. The underlying server infrastructure has proved extremely valuable and continues to automate the collection of OR and ICU data for research and QI purposes.

Team Communication and Situational Awareness

Utilization of telePORT was higher on the paging (request for help) functionality than when being used to exchange information between AAs, suggesting that anesthesiologists appreciated the simple way to request an extra pair of hands. Data also indicated that passive OR monitoring was common. This likely increased the AAs’ situational awareness, which may have preempted additional use of the paging function.

There are concerns about the use of smartphones in the OR, primarily around the potential for spreading infections, distraction for anesthesia team members, and interference with medical equipment [21]; however, they can improve team communication and provide an important learning tool [21]. Apps in the critical care domain are emerging [22], including tools like the Nurse Watch App which provides real-time vital sign monitoring, alarm notifications, and reminders [23], as well as a microblog messaging platform that synchronizes with a patient’s electronic health record and provides a forum for developing and sharing care plans [24].

In the perioperative setting, the VigiVU app (Vanderbilt University) [25] provides anesthesia care providers with notifications of vital sign deviations, changes in patient location,
facilitates team communication and provides high quality video views into each OR. This system was developed and implemented by a much larger team and is more sophisticated than telePORT. To allow other researchers to learn from our lessons and continue development of both telePORT and VitalNode for use in other institutions, their source code has been made freely available (under a Berkeley Software Distribution license) in the LHealth Github repository [26]. Perhaps the biggest accomplishment of the telePORT development process was the facility to live-stream vital signs to other devices for data collection in clinical research and QI projects. VitalNode, similar to the system in use at the University of Michigan [27], has facilitated use of vital signs data in a range of research and quality improvement studies at our institution [28-31].

Implementation Issues: Connectivity and Battery Life

The main feasibility challenge was repeated problems with the hospital’s wireless area network. This led to a decline in use and eventual discontinuation of the project. The signal strength was deemed sufficient when tested, even in the presence of large concrete walls, but the problem with loss of connectivity (or, specifically, the failure to reobtain an internet protocol [IP] address from the dynamic host configuration protocol [DHCP] server) could not be overcome with the hospital network. Our ability to debug the problem was limited and we were unable to ensure that network features that facilitated handover (such as Institute of Electrical and Electronics Engineers 802.11k and 802.11r) were enabled on all access points. It may also have been a limitation of the mobile device we used (the iPod Touch). Had we used a cellular network—enabled device with a data plan we might have used Apple Push Notifications, which has been shown to be a particularly reliable mode of message exchange, even better than regular paging networks [32].

A similar limitation was noted by the developers of VigiVU, who required continued use of traditional pagers as connectivity suffered briefly when switching between certain wireless access points. While seamless handover had improved, it still required an institutional commitment to upgrade and maintain its wireless network infrastructure [27].

During the design stage, there was concern that the iPod Touch battery would not hold sufficient charge to support use during an entire shift. This proved to be a smaller problem than we feared, with batteries draining approximately 9% per hour, enough to last a typical shift. A similar finding was observed with the VigiVU app, in which the battery life left at the end of an 11-hour work day (from full charge) was around 25% [25].

Mobile Apps and Decision Support Systems

The telePORT app extracts data directly from the OR patient monitoring system, but in some institutions, it is possible to use data extracted from an anesthesia information management system (AIMS), which provides for near real-time clinical decision support (CDS). For example, the use of the Smart Anesthesia Manager (SAM, University of Washington, Seattle, USA) [33] demonstrated improved compliance with beta blocker and glucose management protocols [33,34], reduced gaps in blood pressure monitoring [33,35] as well as the duration and frequency of hypotension [36], and even suggested some cost reductions may be feasible [33]. In other cases, intraoperative CDS has shown similar benefits for antibiotic administration and clinical documentation [37]. However, improvements in process measures such as protocol compliance do not necessarily translate into improved outcomes [34,38]. Furthermore, developing such systems presents many challenges, including the need to demonstrate outcome improvements, impact on patient safety, conformance with regulatory requirements [39], and that some improvements may be more easily achieved through the relatively simpler implementation of post hoc reporting [29,40].

Limitations

While the evaluation of the telePORT app was terminated due to an insurmountable technical issue, several shortcomings of the study should be noted prior to any further work on this or a similar initiative. A summative evaluation of the app should have included certain dimensions that we are not able to report here. A structured pre and postimplementation survey of all relevant stakeholders, including the anesthesiologists, may have provided useful opinions on the benefits of the app and process improvements. It might also have been possible to establish a quantifiable measure of an AA’s situational awareness based on their timing on response to critical events. This should include time to provide meaningful assistance as well as arrival, so it would require a careful and agreed upon definition.

We did not obtain a record of the reasons for app messaging (calls for help), including any occurrence of false alarms, which would provide another useful measure of the app’s utility. Finally, while no adverse events related to the use of the app were logged in our institutions’ patient safety and learning system, we also did not implement any measures to track vigilance of the app or resulting clinical errors. Taking this project forward would require a clear definition of the potential points of impact on the existing model of care that can be evaluated postimplementation.

Future Work

Apps like telePORT and VigiVU have the potential to tighten integration with electronic anesthesia records and to support decision support tools, scheduling systems and dashboards to improve perioperative data flow and team communication, with bidirectional data feeds. It should be possible to modify telePORT to use Apple Push Notifications using cellular data to allow better connectivity, and to make it available in the Apple Store, so all anesthesia team members can use it with a (secure) bring-your-own-device model.

The telePORT app was designed for a small group of AAs supporting a larger number of anesthesiologists, but it also has potential as a portable tool for a single anesthesiologist monitoring multiple operating rooms or procedural suites, possibly supporting trainees or other care providers. In either case, an important addition to the tool’s functionality would be the enabling of closed-loop communication via confirmation that a message had been received.

As a next step, our research group has worked closely with pediatric critical care clinicians to develop and evaluate a
preliminary, low-fidelity prototype of the VitalPAD, an app designed to improve the efficiency of clinical decision making, communication, and patient safety in the ICU [41]. The app will ultimately combine information from multiple monitoring and therapeutic devices in a single mobile app, which will include a map overview of the ICU showing clinician assignment, patient status and respiratory support, along with functions for display of patient vital signs, photo-documentation of ABG results, team communication and reminders.

The telePORT app and related projects by other research teams show the potential for innovative, clinically relevant, real-time applications. The research community would benefit greatly from an open protocol for vital signs streaming that would enable new intelligent apps to be rapidly developed and deployed in the hospital setting, with immediate benefits to healthcare professionals, researchers and patients. Protocols developed for the Internet of Things may offer a solution that already has strong community support [11].

Conclusion

We demonstrated that day-to-day activities of the anesthesia team can be supported by a mobile app that integrates data collected in real-time from the OR monitors with a facility for anesthesiologists to send requests for help, as well as team communication features. We overcame significant technical challenges, benefitted from the use of user-centered design, and were able to demonstrate feasibility with the AAs at our institution. Issues with the local wireless network at our site prevented full-scale implementation, but the VitalNode server infrastructure, which collects and stores OR data, continues to benefit ongoing research and QI initiatives.

Acknowledgments

The authors wish to thank Jonathan Hickerson (now EDA Architects) and Jim Agutter at the University of Utah, for their guidance on the design of the app, and the BCCH anesthesia assistants for their enthusiasm for the project, willingness to be observed, contribution to the user-centered design of the interface, and help evaluating the app. This work was supported in part by a Canadian Institutes of Health Research postdoctoral fellowship (MG).

Conflicts of Interest

None declared.

References


Abbreviations

AA: anesthesia assistant
ABG: arterial blood gas
AIMS: anesthesia information management system
BCCH: British Columbia Children’s Hospital
CDS: clinical decision support
CPU: central processing unit
CSV: comma-separated values
CT: computed tomography
DHCP: dynamic host configuration protocol
ECG: electrocardiogram
etCO₂: end-tidal carbon dioxide
HR: heart rate
ICU: intensive care unit
IP: internet protocol
IQR: interquartile range
MRI: magnetic resonance imaging
NIBP: noninvasive blood pressure
OR: operating room
PCAP: packet capture
QI: quality improvement
RR: respiratory rate
SAM: smart anesthesia manager
SpO₂: blood oxygen saturation
telePORT app: Portable Operating Room Tracker app
VoIP: Voice over Internet Protocol
A Postoperative Pain Management Mobile App (Panda) for Children at Home After Discharge: Usability and Feasibility

Dustin Dunsmuir1, MSc; Helen Wu2; Terri Sun1, MD; Nicholas C West1, MSc; Gillian R Lauder1,3, MBBCh; Matthias Görges1,4, PhD; J Mark Ansermino1,3,4, MBBCh

1Department of Anesthesiology, Pharmacology & Therapeutics, University of British Columbia, Vancouver, BC, Canada
2Department of Pharmacology and Therapeutics, McGill University, Montreal, QC, Canada
3Department of Pediatric Anesthesia, BC Children’s Hospital, Vancouver, BC, Canada
4BC Children’s Hospital Research Institute, Vancouver, BC, Canada

Corresponding Author:
Dustin Dunsmuir, MSc
Department of Anesthesiology, Pharmacology & Therapeutics
University of British Columbia
Room V3-353, Clinical Support Building
950 W 28th Ave
Vancouver, BC, VSZ 4H4
Canada
Phone: 1 604 875 2000 ext 5318
Fax: 1 604 875 2668
Email: DDunsmuir@cw.bc.ca

Abstract

Background: Emphasis on outpatient pediatric surgical procedures places the burden of responsibility for postoperative pain management on parents or guardians. Panda is a mobile phone app that provides scheduled medication alerts and allows parents to track their child’s pain and medication administration. We have previously tested and optimized the usability and feasibility of Panda within the hospital setting.

Objective: The purpose of this study was to evaluate and optimize the usability and feasibility of Panda for use at home based on alert response adherence (response to any medication notification within 1 hour) and parents’ satisfaction.

Methods: Parents or guardians of children aged 3 to 18 years undergoing day surgery were recruited to use Panda at home for 1 to 7 days to manage their scheduled medications and to assess their child’s pain. After the surgical procedure, a research assistant guided parents through app setup before independent use at home. We aimed to recruit 10 child-caregiver pairs in each of three rounds of evaluation. Each user’s adherence with the recommended medication alerts was analyzed through audit-trail data generated during the use of the app. We used the Computer System Usability Questionnaire and a poststudy phone interview to evaluate the app’s ease of use and identify major barriers to adoption. Suggestions provided during the interviews were used to improve the app between each round.

Results: Twenty-nine child-caregiver pairs participated in three rounds, using the app for 1 to 5 days. Alert response adherence (response to any medication notification within 1 hour) improved as the study progressed: participants responded to a median 30% (interquartile range [IQR] 22%-33%) of alerts within 1 hour in round 1, and subsequently to median 60% (IQR 44%-64%) in round 2 and median 64% (IQR 56%-72%) in round 3 (P=.005). Similarly, response times decreased from median 131 (IQR 77-158) minutes in round 1 to median 31 (IQR 18-61) minutes in round 2 and median 10 (IQR 2-14) minutes in round 3 (P=.002). Analysis of interview feedback from the first two rounds revealed usability issues, such as complaints of too many pages and trouble hearing app alerts, which were addressed to streamline app function, as well as improve visual appearance and audible alerts.

Conclusions: It is feasible for parents or guardians to use Panda at home to manage their child’s medication schedule and track their pain. Simple modifications to the app’s alert sounds and user interface improved response times.

(JMIR Perioper Med 2019;2(2):e12305) doi:10.2196/12305
KEYWORDS

pain management; pain, postoperative period; outpatients; mobile apps; child; parents

Introduction

Background

The number of pediatric outpatient surgical procedures is increasing [1,2]. A range of medications may be prescribed to manage pain during recovery from ambulatory procedures [3]. Unfortunately, studies have suggested that children frequently experience significant pain following discharge from hospital [4,5] and that poorly managed postsurgical pain can contribute to significant long-term problems [6,7]. Pain may be managed less effectively by parents or caregivers at home than by health care professionals in hospital due to the challenges faced in assessing a child’s pain, following prescription schedules, and calculating doses. There can also be misconceptions about the effects and safety of medications [5,8,9].

There are opportunities for mobile phone apps to guide and improve postoperative pain management for patients after they leave the hospital. Thus far, apps used for self-management of acute sickle cell crisis [10], chronic pediatric cancer pain [11], burn recovery [12], and diabetes [13] have shown promise with patient engagement [10], reduction of anxiety or pain scores [12], improved self-management [13], and overall user satisfaction. A recent review of commercially available mobile phone apps for postoperative pain management demonstrated a lack of evidence-based content [14], and few were designed for use in children. The PediPain app [15] developed by SickKids (a major pediatric hospital in Toronto, ON, Canada) is a pediatric pain and medication dosing app, but it is designed for nurses and doctors in the hospital. Thus, there is still a need for an app to support parents in managing their child’s pain after they leave the hospital.

To address this, the Digital Health Innovation Lab (DHIL) at BC Children’s Hospital (BCCH, Vancouver, BC, Canada) has developed Panda, a mobile phone app designed to support parents in performing three primary tasks: (1) assessing their child’s pain using digitized versions of validated self-report pain tools, (2) scheduling medication reminders using alerts, and (3) tracking medications administered and pain histories. The app has been previously evaluated for safety and ease of use and was improved following several rounds of usability and feasibility testing in hospital [16]. In the feasibility portion, parents used the app while their child was admitted to hospital following surgery. This allowed us to evaluate the app’s usability and feasibility within a controlled environment. The final round of testing demonstrated marked improvement, with 84% (31/37) of app alerts being responded to by parents within 1 hour, and 93% (27/29) of parents reporting that the app was easy to use [16].

The Panda app contains four pediatric self-report pain scales, including established pediatric pain scales, the Faces Pain Scale Revised [17] (ages 4 years and older) and Color Analog Scale [18] (ages 5 years and older), for which we previously demonstrated good agreement between new digital mobile phone versions and the original paper or plastic versions [19].

Additionally, two simplified pain scales, the Simplified Faces Pain Scale and Simplified Concrete Ordinal Scale, have also been designed for mobile phone use and validated in younger children aged 3 to 5 years [20]. Panda was primarily made for parental use, with the exception of the self-reported pain scales, which required the child (aged 3-18 years) to interact with the app to indicate their level of pain.

Objectives

The primary objective of this study was to demonstrate the feasibility of the Panda app with parents or guardians at home; specifically, will parents regularly use the app and consider it a useful aid for administering pain medications to their child? The measure of feasibility was assessed by the degree of alert response adherence (responding to any medication notification within 1 hour) and parental-reported satisfaction with the app during the postoperative care of their child. The secondary objective was to improve the app iteratively by identifying usability issues that arose at home. Usability issues were any undesired or confusing aspects of the app (eg, the appearance or page progression) that caused interactions or behavior that led to undesired or unexpected outcomes, ultimately reducing use of the app and thus negatively affecting feasibility.

Methods

Study Design

This was an observational study to investigate the feasibility of parents using Panda at home to manage their child’s pain after an outpatient surgical procedure. Parents were sent home with the Panda mobile phone app; their use of the app was logged by an automated audit function (recording of all buttons pressed, text entered, and other interactions with the app as time-stamped entries in a log file). Poststudy telephone interviews (Multimedia Appendix 1) were conducted with the participants (parents), and an online questionnaire about the usability of the app was used to collect feedback from each participant. This quantitative usage data and qualitative feedback were then analyzed for common problems and themes and used to improve the app.

Approval from the University of British Columbia-Children’s and Women’s Health Centre of BC Research Ethics Board (H17-00645) was obtained. The study was conducted between May and August 2017 at BCCH in Vancouver, BC, Canada.

Participants

The inclusion criteria for the participants of the study were parent or guardian aged 18 years or older, caring for a child aged 3 to 18 years undergoing an elective ambulatory surgical procedure for which there was minor anticipated postoperative pain. The attending anesthesiologist informed the research assistant (HW) of procedures associated with postoperative pain. The child must have been discharged from hospital on the day of their surgery and have planned postdischarge analgesic medications, including nonsteroidal anti-inflammatory drugs, acetaminophen, or opioids, for at least two days. Parents or guardians or children with hearing or visual impairment.
neurologic injury or psychomotor dysfunction, developmental delay, cognitive neuromuscular inability to operate the app, or inability to follow study instructions in English were excluded. Eligible participants were identified on the operating room schedule by a study research assistant before being approached. Written, informed parental consent and, if appropriate, child assent, were obtained in the surgical day care unit before the child’s operation.

The Panda App
The Panda app has been described previously [16]. The initial app setup includes selecting the appropriate pain scale of the Color Analog Scale, Faces Pain Scale Revised, Simplified Faces Pain Scale, or Simplified Concrete Ordinal Scale (with a recommendation made automatically based on age) and entering a medication schedule according to a child’s prescribed medication. The user has the option of disabling the medication and pain assessment alarms at inconvenient times, such as when they would be sleeping. Then, at the regularly scheduled medication administration times, the Panda app alerts the user via a pop-up notification on their mobile phone. On opening the app, a notification displays the name(s) of the medication(s) to give and offers three response options: proceed, snooze, or skip. Selecting proceed directs the user to the standardized pain scale, and then guides them through medication safety checks before confirming the medication(s) being given to their child. Snooze allows the parent to delay the reminder. Finally, there was an opportunity to skip giving the medication. Additional functions of the app (Figure 1) include editing the patient information and medication schedule, off-schedule medications, off-schedule pain checks, a calendar showing past medication and pain, and a demo mode allowing the user to practice the response to a medication alert.

Figure 1. Home screen of Panda showing the available menu buttons.

Procedures
After written informed consent, the participant completed a prestudy questionnaire about their general knowledge and use of mobile phone apps. For the study, each enrolled participant had the choice of using his or her personal Android or iPhone device, or an iPod Touch provided by DHIL. A research assistant guided the participant in downloading the app and provided an orientation of the Panda app’s primary functions during a suitable time during their hospital stay. The research assistant also guided the participant through the setup of the app by entering their child’s demographics, operation information, and medication schedule as prescribed by the physician after the operation. The research assistant also ensured that the participants were satisfied with the default selected pain scale, which could be changed if requested. Panda includes a tutorial video directly within the app to help the user remember how to use all its functions and features. In addition, each pain scale contains a training page and optional audio instructions for the child.

At home, the participant was expected to use the Panda app for guidance with their child’s pain assessment and management of prescribed analgesic medications. All actions within the Panda app were logged within the audit trail. App data were uploaded to a Research Electronic Data Capture (REDCap) database [21], hosted at the BCCH Research Institute. This enabled the research team to track each participant’s usage of Panda daily during the study period.

The observation period lasted for 1 to 7 days. A research assistant was available by phone, email, or text messaging to answer questions and resolve issues related to the function of the app during regular working hours (9 am - 5 pm, Monday - Friday). The app audit trail on REDCap enabled the research assistant to see when each participant stopped using the app. The research assistant called the participant for the telephone interview within 3 days of their use of Panda stopping completely. The research assistant performed a structured 12-item interview (Multimedia Appendix 1) with each participant about their general satisfaction with the app.
specifically identifying any barriers to use, the utility of specific features, and general ease of use of the app.

A final upload of each participant’s usage data was performed before the app was deactivated and deleted from the participant’s personal device; in the case of a borrowed iPod Touch, it was mailed back to the research office using prepaid envelopes. Finally, participants were emailed an electronic Computer System Usability Questionnaire (CSUQ) [22] survey via REDCap, which consisted of 19 questions inquiring about general usability of the app, each on a 7-point Likert scale. Both the phone interview and the CSUQ measured participant satisfaction. However, the interview was designed specifically for Panda and provided the team with specific qualitative feedback, and the CSUQ provided standardized quantitative data for comparison of rounds. If the survey was not completed within 2 days, the survey email invitation was resent once. If it was still not completed, that participant’s data were analyzed without any CSUQ responses.

Data Analysis
A pragmatic sample size of N=30 was chosen, which is typically suggested for usability evaluations and matches similar studies [11,16,23]. This study was designed using three rounds of data collection, each consisting of 10 participants. After each round, data were extracted from the audit log, including the timing and response (proceed, skip, or snooze) to all app notifications and any off-schedule medications recorded or pain checks done. For each participant, we calculated the duration of study participation, the median response time between each scheduled alert time and the user’s response time, the alert response adherence (response within 1 hour), the median pain scores reported, the proportion of pain checks in which a participant said there was “no pain” (responded to the alert, but specifically said there was no pain and therefore were not shown a pain scale; counted as a score 0), the number of doses of acetaminophen (the most common medication given), and the interval between these doses. These values were compiled across all participants in a round and then compared between rounds using the Kruskal-Wallis test for continuous data; a critical alpha of .05 was used to determine statistical significance. We considered a response to a medication reminder within 1 hour to be compliant with the medication schedule.

Usability was determined by the median CSUQ score, and participant satisfaction was assessed by analyzing the poststudy interview responses. Median values were calculated per question across all completed questionnaires, excluding response of “not applicable” or “N/A.” Responses to interview questions were collated and used to identify common themes [24]. This analysis was done after each round and used to target areas for improvement in app design before the next round. After each round, the software developer (DD) modified the Panda app based on participant feedback collected from the interview and the device-recorded audit of app usage.

Results
Participants
Between May and August 2017, the research assistant identified 68 eligible participants; 40 participants were recruited, and 29 participants successfully completed all questions within the prestudy questionnaire, used the Panda app at home for at least one day (see Figure 2), and then took part in the poststudy telephone interview. Most of the 22 potential participants that declined did not give a reason, but reasons given included the expectation that their child would not be in enough pain to require the app (n=3), they would not require help from an app (n=3), or because they were anxious about their child’s procedure and did not wish to be distracted with participation in the research project (n=1).

Figure 2. Study participant screening and enrollment.
Of the 29 participants, 3 were in their twenties, 11 were in their thirties, 11 were in their forties, and the remaining 4 were older than 50 years. The majority of participants (23/29, 79%) were female, and 90% (26/29) were daily mobile app users. Although they used apps regularly, Panda was a novel app to them as only 52% (15/29) had previously used some form of health or fitness app, and 45% (13/29) reported they would normally use their memory alone to remind themselves when to give medications. In addition, only 38% (11/29) would normally use any type of pain scale (such as numbers 1-10) to have their child communicate their degree of pain. The median age of the participating children was 8 years (interquartile range [IQR] 4-12), and there were 18 males and 11 females. Surgical procedure types included dental (n=6); orthopedic (n=7); general surgery (n=6); ear, nose, and throat (n=6); and plastic surgery (n=4).

**App Usage**

Participants used Panda for a maximum of 5 days during their child’s postoperative recovery at home. Five participants used the DHIL-provided iPod Touch device, four participants used their own Android phones, and the other 20 used their own iPhones. Usage of the app during the study was variable, with some participants reporting that their child was not in enough pain for the app to be needed for very long. The full usage log was obtained from all 29 participants. The duration of app usage varied from only one alert the morning after surgery (one participant in round 2) to 30 alerts over 4 or 5 days (one participant each in rounds 2 and 3). There was no significant difference between the rounds in terms of duration of study participation, median pain scores, proportion of pain scores recorded as “no pain,” number of doses of acetaminophen given, or the interval between these doses (Table 1). Response to alerts improved from one round to the next: participants responded to more alerts within 1 hour ($P=.005$) and median response times decreased significantly ($P=.002$; Table 1 and Figure 3).

Most children reported no pain (score 0/10), low pain (score 1-3/10), or moderate pain (score 4-7/10; see Table 1). However, 6 of 29 (21%) children reported severe pain (score 8-10) at some point during the study, and 4 of 29 (14%) reported severe pain repeatedly (i.e., more than three times). All four of these children were given acetaminophen and ibuprofen, and two of them received an additional analgesic medication, such as oral morphine. The median interval between doses of acetaminophen for these four children were 6 hours, 10 hours, 11 hours, and 19 hours, respectively.

### Table 1. Summary of response times, pain scores, and administration of acetaminophen (N=29).

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Round, median (interquartile range)</th>
<th>Round 1 (n=10)</th>
<th>Round 2 (n=10)</th>
<th>Round 3 (n=9)</th>
<th>$P$ value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time on study (hours)</td>
<td>67 (36-89)</td>
<td>25 (24-92)</td>
<td>51 (27-78)</td>
<td>.39</td>
<td></td>
</tr>
<tr>
<td>Responded to alert before next alert (%)</td>
<td>65 (48-82)</td>
<td>81 (70-100)</td>
<td>89 (75-90)</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>Responded to alert within 1 hour (%)</td>
<td>30 (22-33)</td>
<td>60 (44-64)</td>
<td>64 (56-72)</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Median response time (minutes)</td>
<td>131 (77-158)</td>
<td>31 (18-61)</td>
<td>10 (2-14)</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Number of pain scores recorded</td>
<td>8 (6-13)</td>
<td>8.5 (4-17)</td>
<td>8 (6-12)</td>
<td>&gt;.99</td>
<td></td>
</tr>
<tr>
<td>Median pain score</td>
<td>0 (0-2.5)</td>
<td>3 (0.3-5.6)</td>
<td>0.8 (0-2.5)</td>
<td>.26</td>
<td></td>
</tr>
<tr>
<td>Proportion of pain scores recorded as “no pain” (%)</td>
<td>66 (39-84)</td>
<td>22 (2-58)</td>
<td>20 (17-80)</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Doses of acetaminophen given</td>
<td>4 (2-8)</td>
<td>4.5 (1-7)</td>
<td>5 (2-6)</td>
<td>.92</td>
<td></td>
</tr>
<tr>
<td>Median interval between acetaminophen doses (hours)$^b$</td>
<td>9.3 (4.8-9.9)</td>
<td>11.1 (8.9-13.1)</td>
<td>8.8 (7.1-10.2)</td>
<td>.24</td>
<td></td>
</tr>
</tbody>
</table>

$^aP$ values calculated using the two-tailed Kruskal-Wallis test.

$^b$Prescriptions for acetaminophen varied from every 4 hours (14 participants) to up to every 12 hours (1 participant).
**Participant Feedback**

After the study, 22 of 29 participants (76%) completed the CSUQ. Across three rounds, the rating was median 2 (IQR 1-3) or “agree.” The most positively rated statement was “It was easy to learn using this interface” with a rating of median 1.5 (IQR 1-2). The structured poststudy telephone interview helped identify barriers to using the app. When specifically asked “Was responding to a medication alert clear to you?” 28 of 29 (97%) participants responded positively. After a thematic analysis of participant responses, several issues were identified (Table 2). The first three issues in Table 2 were addressed through changes to the app as described subsequently. As in our earlier study, some issues occurred throughout all three rounds of the study and were not addressed due to conflicting with the purpose of the app (Table 2). For example, one participant stated:

*The main thing is having the medications shift once you give an off-schedule medication.*

Another commented:

*Would prefer faces that children are more familiar with, like emojis. Include a number scale because my son is used to expressing pain on a scale from 1-10.*

There were additional feature requests that were beyond the scope of the app, such as including medication doses and allowing use from two different devices (both parents managing the child’s pain collaboratively).
Table 2. Common feedback from participants by study round (N=29).

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Round 1 (n=10), n</th>
<th>Round 2 (n=10), n</th>
<th>Round 3 (n=9), n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too many app screens during medication safety check</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Trouble hearing alerts</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>App graphics not visually appealing (too “immature”)</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Wanted a more flexible schedule for medications</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Did not want child to use the pediatric pain scales</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 4. The changes made to the medication safety checks after round 1.

App Changes
After the first round, changes were made to reduce the number of pages by combining all medication safety checks into a single page (Figure 4). To address the audibility of alerts, an alert “test” button was added to the initial app setup that triggered an alert 3 seconds later to ensure alerts were both audible and visible on their personal phone. After the second round, the alert sound was changed to a more obvious alert tone, the height and weight were made optional within the demographics, and the home page icons were also redesigned to address the user criticism (Figure 1).

Discussion
Principal Findings
This study evaluated the Panda app in a real-world setting, with parents or guardians successfully using it for their child’s postdischarge pain management at home. There was a statistically significant decrease in the time to respond to the app between rounds of testing and an increased proportion of responses within 1 hour, indicating improved adherence to the app notifications. From interview feedback, key issues with the app were identified and the app was improved between rounds.
According to the prestudy questionnaire, only half of the participants were familiar with health or fitness apps, but the CSUQ results and poststudy interviews suggested that Panda was easy to learn and that the most common action, responding to an alert, was clear to the majority of participants. Common complaints about the app included the graphics and alert sounds, which can be attributed to preference and app aesthetics rather than the app usability and functionality. Nonetheless, we addressed these concerns in an effort to improve Panda’s adoption. Almost all the participants responded positively to Panda and were able to use it with minimal supervision; therefore, it can be said that Panda is feasible for use by parents or guardians at home.

The purpose of this study was to evaluate feasibility and to address outstanding usability issues that may not have been identified during our previous study in the hospital [16], rather than to track any clinical benefits that may be provided by the app. Nonetheless, we noted that 4 of 29 (14%) children who participated in the study did have severe pain that lasted for a significant period of time. There was a high degree of variability in dosing frequency of analgesic medications recorded within the app with the median time for three of these four children being much longer than the prescribed medication schedule despite persistent and severe pain. This may suggest that some parents were not following the clinical guidance provided by the app or that the child was experiencing severe pain despite regular medications; either scenario would warrant clinical attention and provide a window of opportunity for health care intervention to potentially change outcomes. Future development of in-app educational content specific to at-home pain management and two-way communication between the app user and the clinical team may help optimize pain management.

Several research groups are beginning to investigate the use of mobile phone apps to track and promote postoperative rehabilitation, including requirements for patient action within a prescribed recovery program. For example, Jaensson et al [25,26] developed Recovery Assessments by Phone Points (RAPP), a mobile phone app for monitoring and assessing postoperative recovery, and showed that the mobile app was more cost-effective than a control group receiving standard care [27]. Symer et al [28] developed an app to track recovery after abdominal surgical procedures that they suggest will be able to improve outcomes. A recent pilot study by Pecorelli et al [29] reported positive results with a tool to record patient adherence to care processes and patient-reported outcomes as part of an enhanced recovery program for bowel operations. Other researchers have developed an app to monitor wound recovery using mobile phone technology [30,31]. It has also been shown by Lu et al [32] that adherence with pre- and postoperative protocols for reducing surgical site infections can be increased using simple automated messaging systems. These examples illustrate the perceived potential for mobile phone technology in the perioperative period. There are potential advantages for patient outcomes, as well as possible cost-saving benefits for hospitals if these technologies can be shown to facilitate a safe alternative to postoperative care in hospital or reducing returns to hospital for follow-up issues such as pain and wound infection.

Mobile health apps provide an opportunity to optimize pediatric pain management at home [33]. In the context of postoperative care, it may be useful to include educational content on general care, procedure-specific guidance, suggestions for self-management of acute pain, information on prescribed and alternative medications, and instructions for what to do in case of patient deterioration, among other things. These components were beyond the scope of the Panda app, but some are being considered in future iterations ahead of full-scale implementation.

Limitations

This study considered alert response adherence as a surrogate measure of medication adherence, but with the use of the app by parents independently at home there is no way to confirm if the medication was actually given. The app ensured that the participants were made aware of when their child’s medications were due; however, it was the participant’s (and patient’s) choice whether to comply or not.

The target primary user for Panda in this study was exclusively a parent or guardian. Despite the child age range including adolescents, we previously found teens to have a low task completion rate within Panda [16]. A pain management app for teens should be designed specifically for them, and any additional educational content would likely be presented quite differently than for parents or guardians. For example, Stinson et al [11] gamified pain management to motivate adolescents with cancer to use their app. Our study did not collect any feedback directly from the pediatric patients given their varying age but instead relied on feedback from the main app user, the parent. Feedback on the pain scales was previously collected in our testing of the digital versions of the scales [19].

The heterogeneity of the study cohort meant that some patients reported very little pain during their postoperative recovery. This may have biased their parents against using the app, which is hard to interpret from the data available. Future evaluation of the app may be targeted at more specific age and procedure groups when we aim to evaluate the impact of Panda on patient outcomes.

Although some users expressed that the app was easy to use, there were additional features suggested, which were not always aligned with the current goals of the app (eg, improving simplicity over ensuring medications were administered safely). The main goal of the Panda app is to help with management of postoperative pain by increasing adherence to the medication schedule prescribed or recommended by their health care providers. Panda allowed participants to determine the time of the first dose of medication and to turn off any alert (such as alerts while sleeping), but all alert times were predetermined according to the selected time interval (give medication every 4 hours, 6 hours, etc). Participants expressed a desire for more control of the schedule by being able to adjust the predetermined schedule to their convenience; however, allowing this would require an intelligent algorithm to ensure that the medication schedule remained in accordance with safe dosing. The addition of dynamic pain medication scheduling to the app may hinder the simplicity and predictability of the schedule.
Future Work
We are planning to enhance Panda’s utility through the addition of two-way communication features before releasing it on the public Google and iOS app stores. Clinicians have expressed the usefulness of better follow-up information on their patients postoperatively. It is clear from some of the children’s pain scores that medication alerts alone may not be enough to eliminate severe pain during the recovery period, and intelligently communicating these pain scores to clinicians could identify specific children who might require additional interventions. Additionally, parents would benefit from a new and more convenient means of communication with the clinical team during their child’s postoperative recovery. Hence, we will be adding a two-way communication system to the Panda platform, allowing the clinicians access to a dashboard view of their patients’ pain scores and medication usage as sent from the Panda app.

Conclusions
The Panda app is feasible to be used by parents at home for the management of their child’s postoperative pain. Apps such as Panda can take advantage of the ubiquity of mobile phones to provide a useful pain management tool in parents’ pockets. Overall, the Panda app can serve as a support system for parents to help manage their child’s pain after discharge from hospital, potentially preventing unnecessary pain and associated sequelae and the need to reaccess health care resources.

Acknowledgments
The authors would like to thank all participating parents, as well as the nursing staff in the Surgical Day Care Unit at BCCH.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Telephone interview questions.

References


Abbreviations

BCCH: BC Children’s Hospital
CSUQ: Computer System Usability Questionnaire
DHIL: Digital Health Innovation Lab
IQR: interquartile range
REDCap: Research Electronic Data Capture

© Dustin Dunsmuir, Helen Wu, Terri Sun, Nicholas C West, Gillian R Lauder, Matthias Görges, J Mark Ansermino. Originally published in JMIR Perioperative Medicine (http://periop.jmir.org), 04.07.2019. 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 Perioperative Medicine, is properly cited. The complete bibliographic information, a link to the original publication on http://periop.jmir.org, as well as this copyright and license information must be included.
Comparing the Relaxing Effects of Different Virtual Reality Environments in the Intensive Care Unit: Observational Study

Stephan M Gerber¹, MSc; Marie-Madlen Jeitziner², PhD; Simon D Sänger¹, MSc; Samuel E J Knobel¹, MSc; Laura Marchal-Crespo¹, Prof Dr; René M Müri¹,², Prof Dr; Joerg C Scheffold³, Prof Dr; Stephan M Jakob², Prof Dr; Tobias Nef⁴, Prof Dr

¹Gerontechnology & Rehabilitation Group, University of Bern, Bern, Switzerland
²Department of Intensive Care Medicine, University Hospital Bern (Inselspital), University of Bern, Bern, Switzerland
³Department of Neurology, University Neurorehabilitation, University Hospital Bern (Inselspital), University of Bern, Bern, Switzerland
⁴ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

Corresponding Author:
Tobias Nef, Prof Dr
Gerontechnology & Rehabilitation Group
University of Bern
Murtenstrasse 50
Bern, 3008
Switzerland
Phone: 41 316327579
Email: tobias.nef@artorg.unibe.ch

Abstract

Background: After a prolonged intensive care unit (ICU) stay, approximately 50%-75% of all critically ill patients suffer from neurocognitive late effects and a reduction of health-related quality of life. It is assumed that the noisy and stressful ICU environment leads to sensory overload and deprivation and potentially to long-term cognitive impairment.

Objective: In this study, we investigated three different virtual reality environments and their potentially restorative and relaxing effects for reducing sensory overload and deprivation in the ICU.

Methods: A total of 45 healthy subjects were exposed to three different environments, each 10 minutes in length (dynamic, virtual, natural, and urban environments presented inside the head-mounted display, and a neutral video on an ICU TV screen). During the study, data was collected by validated questionnaires (ie, restoration and sickness) and sensors to record physiological parameters (240 hertz).

Results: The results showed that the natural environment had the highest positive and restorative effect on the physiological and psychological state of healthy subjects, followed by the urban environment and the ICU TV screen.

Conclusions: Overall, virtual reality stimulation with head-mounted display using a dynamic, virtual and natural environment has the potential, if directly used in the ICU, to reduce sensory overload and deprivation in critically ill patients and thus to prevent neurocognitive late effects.


KEYWORDS
virtual reality; critical illness; intensive care unit; neurocognitive late effects; nature; urban; stimulation

Introduction

Approximately 50%-75% of critically ill patients experiencing a prolonged intensive care unit (ICU) stay will suffer from neurocognitive late effects [1,2]. Neurocognitive late effects (eg, deficits in learning and memory, slowed information processing, attention, concentration, and executive function) of critically ill patients have been associated with a reduced ability to cope with everyday activities and a significant reduction in quality of life after discharge [2-5]. In addition to the medical condition of critically ill patients itself (eg, organ failure, delirium, and muscle weakness and atrophy), the stressful environment is also an important contributor to neurocognitive late effects [6,7]. ICU patients are exposed to high noise levels, artificial light, and isolation from their common environment,
which leads to sensory overload and deprivation and potentially to long-term cognitive impairment [7-9]. There is evidence that the early introduction of an intervention in the ICU has the greatest impact on long-term outcome [1,10]. Current clinical practice in the ICU to reduce functional impairment includes early mobilization and physiotherapy, but there is less consensus about how to reduce or prevent neurocognitive late effects [11].

We propose using virtual reality (VR) technology with head-mounted display to comfort and reduce the stress of ICU patients. A state-of-the-art, head-mounted display for the presentation of stereoscopic images and three-dimensional sounds (developed for the gaming industry) provides sensors for measuring head movements, which allows the user to look around in the virtual environment. Thus, wearing a head-mounted display enables patients to experience virtual environments with their visual and auditory senses while being highly immersed [12]. Since vision and hearing are completely controlled by the head-mounted display, patients do not notice anything from their surroundings, thus helping them avoid the noisy and stressful ICU environment. The virtual environment presented inside a head-mounted display is a computer-generated, visual, and auditory replication of a real (eg, a forest or city) or imaginary (eg, a futuristic city) environment. Virtual environments must be designed carefully to avoid negative side effects (eg, cybersickness, oculomotor problems) [13,14]. This leads us to the question of what virtual environment would comfort ICU patients the most.

In a previous study, we showed that VR exposure to a virtual environment that replicates beautiful nature scenes is feasible and potentially beneficial for critically ill patients. The stimulation had a strong relaxing effect without any adverse effects. Participants reported that the VR exposure was calming and that they were immersed in the virtual environment [15]. Our head-mounted display complied with hygiene requirements in the ICU, and patients could use the head-mounted displays while lying in bed. Based on these preliminary results, we believe that head-mounted displays can be used as an early rehabilitative intervention to provide cognitive stimulation with the goal of reducing neurocognitive late effects.

In the present study, we compare the relaxing effects of a virtual nature environment with a virtual urban environment, in combination with noise-cancelling headphones. As a control condition, we selected watching a movie about intuition on a TV screen (movie), with each interaction lasting 10 minutes. In the control condition, we selected watching a neutral movie on a predefined path. In the control condition, we selected watching a neutral movie on a predefined path. The study was conducted in a one-bed cubicle in the ICU of the University Hospital Bern. To stimulate the subjects, either a head-mounted display (HTC Vive, High Tech Computer Corporation, Taoyuan, Taiwan) or a classical ICU TV screen (Meditec TV, Bewatec, Telgte, Germany) was used. The head-mounted display (HTC Vive, High Tech Computer Corporation, Taoyuan, Taiwan) was used. The head-mounted display had a resolution of 1080 × 1200 pixels per eye, a field of view of 110 degrees, and a refresh rate of 90 hertz (Hz), whereas the ICU TV screen had an 11.5-inch display and an aspect ratio of 4:3.

Subjects were stimulated two times with the head-mounted display (ie, virtual nature and urban VR stimulation, Figure 1) and once with the gold standard (control condition), a classical ICU TV screen (movie), with each interaction lasting 10 minutes. The nature VR stimulation consisted of a large island surrounded by water, with green hill areas, beaches, forests, plantations, and several animals such as elephants, giraffes, dolphins, birds, and butterflies. The designed urban VR stimulation consisted of a busy downtown, followed by a more relaxed old town. Pedestrians, cars, birds, aircrafts, and construction sites were among the many objects that could be found in the environment. In both cases, subjects were walking on a predefined path. In the control condition, a neutral documentary movie about intuition was presented (ie, combination of nature and urban scenes).

Physiological parameters (ie, noninvasive arterial blood pressure, heart rate, and respiratory rate) were monitored by the

Methods

Participant Recruitment and Demographics

The study was conducted following the current version of the Declaration of Helsinki. All subjects participating in the study were recruited through the University of Bern, and the study was approved by the Ethics Committee of the Canton of Bern, Switzerland (KEK-Nr. 2017-02195). All subjects signed written informed consent before inclusion. Furthermore, specific consent was obtained to publish identifying information and images in an online open-access publication. The main exclusion criteria were auditory and visual impairments and being below the age of 18.

Neurocognitive Stimulation

The study was conducted in a one-bed cubicle in the ICU of the University Hospital Bern. To stimulate the subjects, either a head-mounted display (HTC Vive, High Tech Computer Corporation, Taoyuan, Taiwan) or a classical ICU TV screen (MediTec TV, Bewatec, Telgte, Germany) was used. The head-mounted display had a resolution of 1080 × 1200 pixels per eye, a field of view of 110 degrees, and a refresh rate of 90 hertz (Hz), whereas the ICU TV screen had an 11.5-inch display and an aspect ratio of 4:3.
in-house monitor system of the University Hospital (Carescape Monitor B650, GE Healthcare, Little Chalfont, United Kingdom). Respiration and heart rate were measured at a frequency of 240 Hz, whereas blood pressure was measured every second minute. To simulate a real ICU scenario, the physiological monitor system constantly produced alarm sounds (e.g., simulated arrhythmia). In addition, a noise meter was set up next to the subject to monitor and guarantee the overall noise level in the room was above 50 decibels. The bed was tilted up so that the subjects were in an upright position. The system was approved by the University Hospital’s medical-technical department, except for use in critically ill patients and subjects (i.e., hygiene and medical eligibility approved). To evaluate the perceived restorativeness of each stimulation, the PRS-11[23] (measured by a seven-point scale from not at all to completely) and for discomfort the SSQ [24] (measured by a four-point scale, from none to severe) were used. Both score scales were normalized on a score scale between zero and one.

**Figure 1.** Left: VR urban environment; right: VR nature environment. VR: virtual reality.

### Study Design and Procedure

To account for the influence of the order of the stimulations, the subjects were randomly assigned in a pseudo cross-validation to one of the three groups with different stimulation sequences. First, subjects were instructed and prepared (i.e., lying on the bed, **Figure 2**) for the experiment, followed by a recovery phase of 10 minutes to stabilize physiological parameters at subjects’ baseline. Next, subjects were assigned randomly to one of the three groups and stimulated with the three environments with a break of approximately 10 minutes in between. After each stimulation, they were asked to fill out a questionnaire of the perceived restorativeness and discomfort. Before the stimulation, subjects filled out a questionnaire about demographics. All instructions were given prior to the stimulation, and participants could move their head to explore the environment while seated. For each participant, the experiment lasted approximately 80 minutes.
Statistical Analysis

To analyze the calming and relaxing effect (ie, negative time effect is defined as the reduction of physiological parameters during the course of the stimulation) of an environment, the correlation between time and the physiological measurement was calculated and tested on their significance using a one-tailed $t$ test against zero. To analyze the differences in questionnaire ratings between the stimulations, one-way analysis of variance (ANOVA) was used, followed by a post hoc analysis using paired $t$ tests and Bonferroni correction for $P$ value adjustment. The analysis of the physiological parameters and the questionnaires was conducted by using R for statistics (The R Foundation, Vienna, Austria) and MATLAB 2018b (MathWorks, Natick, USA).

Website

A video of the two VR stimulations and the movie about intuition can be found on our website [25].

Data and Code Availability

All relevant codes and data supporting the findings are presented in this paper. Further code and data of the study are available upon request.

Results

Perceived Restoration

A total of 45 (22 male and 23 female) adults between the ages of 22 and 87 (mean 59; SD 16) participated in the study. Overall, 29 participants grew up in the countryside and 16 in an urban environment, and they were all in good health. Analysis using ANOVA showed a significant difference between the three stimulation environments ($F_{2}=15.86; P<.001$). The nature VR stimulation had the highest perceived restoration (mean 0.773; SD 0.142) and was closest to the maximum of the score scale (with none=0 and high=1), followed by urban VR (mean 0.65715; SD 0.187) and ICU TV stimulation (mean 0.5854; SD 0.136). A post hoc analysis showed a significant difference between nature VR and both urban VR ($P=.002$) and ICU TV ($P<.001$), but not between urban VR and ICU TV stimulation ($P=.12$).

In all three environments, nausea, oculomotor problems, and disorientation were close to the minimum of the score scale (Figure 3). The only significant difference in discomfort between the environments was found in disorientation ($F_{2}=3.67; P=.03$).
**Physiological Measurements**

At the beginning of the recovery phase, respiration rate was an average of 19.35 Breaths/minute (SD 5.08), heart rate was an average of 69.88 Beats/minute (SD 11.03), and the mean blood pressure was 94.3 mmHg (SD 11.81). As shown in Figure 4, nature and urban VR stimulation were associated with a reduction in respiration rate, whereas during ICU TV stimulation respiration rate increased. The negative correlation of respiration rate and time (negative time effect) was highest and only significant in nature VR stimulation (95% CI –Inf to –0.024; $t_{42}=-2.13; P=.02$), but neither in urban VR stimulation (95% CI –Inf to 0.052; $t_{40}=-0.69; P=.25$) nor in ICU TV stimulation (95% CI –Inf to 0.132; $t_{41}=1.31; P=.90$). With respect to heart rate, only nature VR decreased during the course of the stimulation, whereas in urban VR and ICU TV stimulation heart rate remained more or less constant. However, there was a significant negative time effect in heart rate in nature VR stimulation (95% CI –Inf to –0.177; $t_{2}=-7.53; P<.001$), in urban VR stimulation (95% CI –Inf to 0.161; $t_{40}=-7.35; P<.001$) and in ICU TV stimulation (95% CI –Inf to 0.141; $t_{38}=1.31; P<.001$). Furthermore, the mean blood pressure did not increase or decrease during any of the three stimulations.
Discussion

Primary Results

In this study, we investigated the subjective restorative and relaxing effect of three different stimulation methods (dynamic, virtual, nature and urban environments presented inside a head-mounted display and a neutral video on an ICU TV screen) in healthy subjects. In line with our hypothesis, there was a significantly higher restorative and relaxing effect in the virtual nature environment compared to the virtual urban environment and the ICU TV screen. Additionally, in all cases, the three different methods evoked none to minor negative reactions.

Perceived Restoration and Comfort

The main finding in line with the attention restoration theory was that the highest perceived restorative effect was found in virtual nature stimulation, followed by the virtual urban stimulation and the movie presented on the ICU TV screen [19,26]. Furthermore, the three different stimulation methods showed none to minor negative side effects, like nausea, oculomotor problems, or disorientation. The slightly higher discomfort in the virtual urban environment was possibly due to unexpected 90 degree turns (ie, when walking around a building corner), whereas in the virtual nature environment no turns were present.

Physiological Measurements

The second main finding was that the nature VR stimulation with head-mounted display revealed the highest and only significant reduction in respiration rate. In case of heart rate, only the nature VR stimulation showed a relaxing effect, whereas in mean blood pressure no effect was found. Overall, physiological parameters showed that nature VR stimulation by using head-mounted displays revealed the greatest effect on the parasympathetic nervous system and thus was in line with the stress reduction theory suggestion that natural environments are the most restorative environment [15,16,27].

The effect on the parasympathetic nervous system may not result just from the natural environment, but may also be influenced by the head-mounted display and the noise-cancelling headphones, which helped to reduce the sensory overload (eg, alarms, noise) and deprivation (eg, missing daylight) in the ICU.
One may argue that this effect should also be present in the urban environment, but there is a significant difference between the stimuli of these environments. While the sound of the natural environment mostly consisted of crashing waves and animal sounds, the soundscape of the urban scene was overloaded with typical city sounds [28]. These sounds may be as irrelevant and annoying as the noises in the ICU and therefore also cause sensory overload. Furthermore, due to memories and past experiences in the different environments, not all participants will respond the same to the different stimuli. Therefore, if in the past participants had a bad experience with certain stimuli inside a natural (eg, animals) or an urban environment (eg, car accident) it may evoke bad memories and thus influence physiological parameters negatively, with the opposite being true for positive experiences.

Limitations

One of the limitations of this study is that all subjects were healthy subjects, so it is unclear whether the results could be transferred to critically ill patients. Another limitation was that cultural, living environment, and educational differences between the participants were recorded as a covariate but were not considered. Subjects from a rural origin may react differently to the stimulation and have different preferences than a subject who was born and grew up in a city. Furthermore, the different movement patterns inside the environments might have an influence on the perceived restorativeness.

Conclusion

Overall, three environments that were different in technical setup and design were tested safely and successfully. We showed that VR stimulation by using a head-mounted display in combination with a dynamic, virtual, natural environment had the highest restorative effect on the physiological and psychological state of a subject. Furthermore, it was shown that the effect was not simply due to the isolation from the stressful environment but actually due to the composition of the natural environment. Therefore, the method of VR nature stimulation by using a head-mounted display may have potential as an early-intervention method directly used in the ICU to reduce sensory overload and deprivation and thus prevent neurocognitive late effects.

Acknowledgments

The authors thank our clinical doctor and nursing team from the Department of Intensive Care Medicine at University Hospital Bern (Inselspital) for their commitment and helping to test the subjects. Furthermore, we would like to show our gratitude to all co-workers at the Gerontechnology and Rehabilitation Research Group at University of Bern for ideas and support during the study. The study was funded by the University of Bern the Gerontechnology and Rehabilitation Research Group at ARTORG Center for biomedical engineering and the University Hospital Bern (Inselspital) Department of Intensive Care Medicine.

Authors' Contributions

SMG, MMJ, RMM, SMJ, and TN designed the study. SMG and SDS developed the setup and measured the participants together with MMJ. SMG, SDS, SEJK, and JCS analyzed the data. SMG, MMJ, SEJK, LM, RMM, JCS, SMJ, and TN wrote the manuscript. All authors approved the final manuscript. TN and SMJ share last authorship.

Conflicts of Interest

MMJ and JCS report grants from Orion Pharma, Abbott Nutrition International, B Braun Medical AG, CSEM AG, Edwards Lifesciences Services GmbH, Kenta Biotech Ltd, Maquet Critical Care AB, Omnicare Clinical Research AG, Nestle, Pierre Fabre Pharma AG, Pfizer, Bard Medica SA, Abbott AG, Anandic Medical Systems, Pan Gas AG Healthcare, Bracco, Hamilton Medical AG, Fresenius Kabi, Getinge Group Maquet AG, Dräger AG, Teleflex Medical GmbH, Glaxo Smith Kline, Merck Sharp and Dohme AG, Eli Lilly and Company, Baxter, Astellas, Astra Zeneca, CSL Behring, Novartis, Covidien, and Nycomed outside the submitted work. The money was paid into departmental funds, and there was no personal financial gain.

References


Abbreviations

ANOVA: analysis of variance
Hz: hertz
ICU: intensive care unit
Inf: infinity
PRS-11: Perceived Restorativeness Questionnaire
SSQ: Simulator Sickness Questionnaire
VR: virtual reality
Original Paper

Impact of an Intensive Care Information System on the Length of Stay of Surgical Intensive Care Unit Patients: Observational Study

Camille Havel¹, MD; Jean Selim¹,², MD; Emmanuel Besnier¹,², MD; Philippe Gouin¹, MD; Benoit Veber¹, PhD; Thomas Clavier¹,², MD

¹Department of Anesthesiology and Critical Care, Rouen University Hospital, Rouen, France
²Normandie Univ, UNIROUEN, INSERM U1096, Rouen, France

Corresponding Author:
Camille Havel, MD
Department of Anesthesiology and Critical Care
Rouen University Hospital
1 rue de Germont
Rouen, 76000
France
Phone: 33 2 32 88 89 90
Email: camille.havel@chu-rouen.fr

Abstract

Background: The implementation of computerized monitoring and prescription systems in intensive care has proven to be reliable in reducing the rate of medical error and increasing patient care time. They also showed a benefit in reducing the length of stay in the intensive care unit (ICU). However, this benefit has been poorly studied, with conflicting results.

Objective: This study aimed to show the impact of computerization on the length of stay in ICUs.

Methods: This was a before-after retrospective observational study. All patients admitted in the surgical ICU at the Rouen University Hospital were included, from June 1, 2015, to June 1, 2016, for the before period and from August 1, 2016, to August 1, 2017, for the after period. The data were extracted from the hospitalization report and included the following: epidemiological data (age, sex, weight, height, and body mass index), reason for ICU admission, severity score at admission, length of stay and mortality in ICU, mortality in hospital, use of life support during the stay, and ICU readmission during the same hospital stay. The consumption of antibiotics, biological analyses, and the number of chest x-rays during the stay were also analyzed.

Results: A total of 1600 patients were included: 839 in the before period and 761 in the after period. Only the severity score Simplified Acute Physiology Score II was significantly higher in the postcomputerization period (38 [SD 20] vs 40 [SD 21]; P<.05). There was no significant difference in terms of length of stay in ICU, mortality, or readmission during the stay. There was a significant increase in the volume of prescribed biological analyses (5416 [5192-5956] biological exams prescribed in the period before Intellispace Critical Care and Anesthesia [ICCA] vs 6374 [6013-6986] biological exams prescribed in the period after ICCA; P=.002), with an increase in the total cost of biological analyses, to the detriment of hematological and biochemical blood tests. There was also a trend toward reduction in the average number of chest x-rays, but this was not significant (0.55 [SD 0.39] chest x-rays per day per patient before computerization vs 0.51 [SD 0.37] chest x-rays per day per patient after computerization; P=.05). On the other hand, there was a decrease in antibiotic prescribing in terms of cost per patient after the implementation of computerization (€149.50 [$164 USD] per patient before computerization vs €105.40 [$155 USD] per patient after computerization).

Conclusions: Implementation of an intensive care information system at the Rouen University Hospital in June 2016 did not have an impact on reducing the length of stay.


KEYWORDS

intensive care unit; length of stay; software; critically ill patient
**Introduction**

**Background**

Medical and paramedical staff working in intensive care units (ICUs) need to collect and process a lot of data regarding patient care (eg, vital parameters, drug prescriptions and biological analyses, and changes in multiple daily prescriptions). The management of these data represents a time that is not directly devoted to patient care, and this can lead to errors in the prescription or delivery of treatments or may delay care.

To increase both the time dedicated to patients and the quality of care, ICU has been equipped with clinical information system (intensive care information system, ICIS) [1-5] over the past 20 years. These clinical information systems have demonstrated their effectiveness in reducing medical errors and improving patient safety [6]. Indeed, many studies have focused on showing the effect of computerization on the reduction of medical errors, whether in ward or in ICU. They all showed a reduction in the rate of medical errors by up to 80% and a better detection of errors before they were committed [7-10]. In addition, it has been shown that errors made were less serious and had fewer serious side effects [10]. It has also been shown that clinical information systems lead to an increase in the amount of time dedicated to patients’ care through a reduction in the time spent consulting and collecting the medical data [11,12].

The impact of information systems on the departments’ organization and the length of stay has been less studied. It is known that adverse drug events in hospitalized patients increase the length of stay and induce extra cost [13]. Thus, given the beneficial impact of information systems on medical and treatment delivery errors, we could have expected a beneficial impact of these systems on the length of hospital stay. However, the impact of computerization on the length of stay in ICU has been poorly studied, with conflicting data. In 2014, Levesque et al found a reduction in the length of stay in ICU of approximately 2 days following ICIS implementation [14], but this study only concerned critically ill patients from a specialized ICU. In addition, a study conducted on all the services of a university hospital before and after computerization found an overall reduction in the length of stay in the medical and surgical services but an increase in the length of stay in most ICUs [15]. Finally, a third study did not find any impact of ICIS on length of stay in ICU [16].

Intellispace Critical Care and Anesthesia (ICCA) is a recent ICIS developed by Philips Healthcare (Amsterdam, Holland). It allows a precise prescription of treatments (administration route, doses, administration times, and duration) as well as the automatic collection of a patient’s vital constants from bedside monitors and ventilators. It also contains the medical file and gathers the information essential to the patient’s care. The literature on the impact of ICIS on the length of stay is controversial, and there is no report on the impact of ICCA software (Philipps) on the quality of care for critically ill patients.

**Objective**

The objective of this study was to assess the potential impact of ICCA on the length of stay of patients admitted in ICU.

**Methods**

**Study Design**

This was a retrospective before-after observational study conducted in an adult surgical ICU of a tertiary care hospital. This study received a favorable opinion from the Ethics Committee for Non-Interventional Research of the Rouen University Hospital (nE2018-55). We compared 2 periods: 1 period before the implementation of ICCA (from June 1, 2015, to June 1, 2016) and the other after its implementation (from August 1, 2016, to August 1, 2017). We excluded the 2 months following implementation of ICCA from the analysis to avoid the bias of software discovery by medical and paramedical staff, as previously described [14].

During these 2 periods, all adult patients hospitalized in ICU were included. For each patient, we collected epidemiological data (age, sex, weight, height, and body mass index), reason for ICU admission (trauma, vascular surgery, visceral surgery, thoracic surgery, other major surgery, medical or surgical sepsis, posttraumatic and postoperative hemorrhagic shock, and other reasons), severity score at admission (Simplified Acute Physiology Score II [SAPSII]), mortality and length of stay in ICU, readmission in ICU, and the use of life support during the stay (catecholamine, mechanical ventilation, and dialysis). Concerning the readmission rate, we have considered any readmission in ICU during the same hospital stay, whether or not it was for the same reason; only the first stay was analyzed.

The data collection was based on the patient’s hospitalization report available on the hospital’s software.

During these periods, we also collected the number of chest x-rays prescribed during the stay; the consumption of biological analyses (bacteriological, biochemical, hematological, pharmacological, and virological analyses), collected in terms of total cost and prescribed volume month by month; and antibiotic consumption (in terms of total cost and delivery volume from the hospital’s central pharmacy). The consumption of complementary exams (chest x-rays and biological analyses) as well as antibiotic consumption only concerned the ICU stay included in the analysis if the patient had been readmitted in ICU. Concerning biological analyses and antibiotic consumption, we have obtained these data from persons not working in our department; their data extracting method is unknown to us. It should be noted that during the 2 years of inclusion of our study, the recommendations of the French Society of Anesthesia and Resuscitation regarding the correct prescribing practices for complementary exams were not modified, nor were the department’s internal recommendations.

The primary objective of this study was to compare the average length of stay in surgical ICU before and after implementation of ICCA. The secondary objectives were to analyze mortality, readmission in ICU during the same stay, as well as the consumption of additional examinations and antibiotic therapy (in terms of total quantity, total cost, and cost per patient over...
the period studied) as indirect indicators of the quality of medical prescriptions.

**Intellispace Critical Care and Anesthesia**

The computerization of our ICU took place in June 2016. It was based on the ICCA software implementation. This computer support includes the patient’s medical file, computerized medical prescriptions, as well as the patient’s monitoring data. Computerized prescription is done using drop-down menus, classified by category (continuous, discontinuous drugs, additional tests, and biological analyses). All treatments referenced at the hospital central pharmacy are integrated into the software, and there are preconfigured prescriptions with proposed treatment regimens for some commonly used treatments (Multimedia Appendix 1). Similarly, most of the additional tests are available as drop-down menus with ticked items, such as the additional biological tests available at the hospital (Multimedia Appendix 1).

The patient monitoring signs include the different vital constants, ventilation parameters, various drainage systems, and so on, which are necessary for monitoring critically ill patients (Multimedia Appendix 2). Some data are extracted from the ICU devices (data from scopes, ventilators, and electric syringes), allowing real-time monitoring of the patient’s progress as well as the various therapies administered. There is also the possibility of manually entering the data, when they are not automatically collected (eg, diuresis), when the automatic data are incorrect, or when the feedback is interrupted (eg, connection broken between the ICU devices and the software).

**Statistical Analysis**

On the basis of the publication of Levesque et al and our estimated mean length of stay in ICU (7 days), we assumed that a difference of 2 days (with a SD of 12 days) between the 2 groups would be clinically significant [14]. On the basis of these results, assuming that the SD was the same between populations and using a power of 0.90 with a statistical significance level of .05, we estimated that a minimum of 757 patients should be analyzed in each group, representing an inclusion duration of 1 year for each period.

Patients from both periods were compared by statistical analyses based on comparisons of means (Student test and Mann-Whitney test) and contingency analyses (chi-square). The statistical analyses were performed using the GraphPad Prism 7 software (GraphPad Software, USA). Quantitative variables are presented as mean (SD) if the distribution respected the normal law or median (interquartile range); if not, qualitative variables are expressed as absolute number and percentage.

**Results**

**Epidemiological Data**

Over the study periods, 1600 patients were included: 839 during the before period and 761 during the after period. The 2 groups of patients were comparable in terms of age and sex (Table 1). We observed a significant difference concerning the reason for ICU admission with more bleeding shock after computerization (6.4% (54/839) before ICCA vs 9.3% (71/761) after ICCA; \( P=.04 \)) and less other major surgery after computerization (10.9% (91/761) before ICCA vs 6.6% (50/761) after ICCA; \( P=.003 \)). SAPSII at ICU admission was lower in the before group (SAPSII: 38 [SD 20] vs 40 [SD 21] after computerization; \( P<.05 \); Table 1).
Table 1. Epidemiological data.

<table>
<thead>
<tr>
<th>Data analyzed</th>
<th>Before ICCA</th>
<th>After ICCA</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>57.8 (17.8)</td>
<td>56.7 (18.8)</td>
<td>.22</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>537 (64.0)</td>
<td>505 (66.4)</td>
<td>.34</td>
</tr>
<tr>
<td>Female</td>
<td>302 (36.0)</td>
<td>257 (33.6)</td>
<td>.34</td>
</tr>
<tr>
<td>Body mass index, mean (SD)</td>
<td>26.6 (6.3)</td>
<td>26.3 (6)</td>
<td>.38</td>
</tr>
<tr>
<td>Simplified Acute Physiology Score II, mean (SD)</td>
<td>38 (20)</td>
<td>40 (21)</td>
<td>.02</td>
</tr>
</tbody>
</table>

Reason for hospitalization, n (%)

<table>
<thead>
<tr>
<th>Reason for hospitalization</th>
<th>Before ICCA</th>
<th>After ICCA</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sepsis</td>
<td>114 (13.6)</td>
<td>106 (13.9)</td>
<td>.88</td>
</tr>
<tr>
<td>Trauma</td>
<td>132 (15.8)</td>
<td>145 (19.0)</td>
<td>.09</td>
</tr>
<tr>
<td>Vascular surgery</td>
<td>64 (7.6)</td>
<td>57 (7.5)</td>
<td>.92</td>
</tr>
<tr>
<td>Gastrointestinal surgery</td>
<td>97 (11.6)</td>
<td>78 (10.2)</td>
<td>.42</td>
</tr>
<tr>
<td>Bleeding shockd</td>
<td>54 (6.4)</td>
<td>71 (9.3)</td>
<td>.04</td>
</tr>
<tr>
<td>Thoracic surgery</td>
<td>41 (4.9)</td>
<td>36 (4.7)</td>
<td>.91</td>
</tr>
<tr>
<td>Other major surgery</td>
<td>91 (10.9)</td>
<td>50 (6.6)</td>
<td>.003</td>
</tr>
<tr>
<td>Other cause</td>
<td>244 (29.1)</td>
<td>218 (28.7)</td>
<td>.83</td>
</tr>
</tbody>
</table>

aICCA: Intellispace Critical Care and Anesthesia.
bSignificant results.
cAll medical or surgical sepsis.
dTraumatic or postoperative bleeding shock.

Length of Stay and Mortality

All the ICU stay characteristics of the 2 groups are summarized in Table 2. There was no difference in the length of stay in ICU between the groups (7.0 (SD 9.3) days before computerization vs 7.4 (SD 9.9) days after computerization; P=.37). Mortality in the 2 groups did not differ significantly, nor did the readmission rate during the same stay and the use of different life supports (Table 2).

Table 2. Intensive care unit stay characteristics of the 2 groups.

<table>
<thead>
<tr>
<th>Data analyzed</th>
<th>Before ICCA</th>
<th>After ICCA</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of stay (days), mean (SD)</td>
<td>7.0 (9.3)</td>
<td>7.4 (9.9)</td>
<td>.37</td>
</tr>
<tr>
<td>Mortality, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In intensive care unit</td>
<td>139 (16.6)</td>
<td>128 (16.8)</td>
<td>.89</td>
</tr>
<tr>
<td>In hospital</td>
<td>181 (21.6)</td>
<td>171 (22.5)</td>
<td>.67</td>
</tr>
<tr>
<td>Re-admission during the same stay, n (%)</td>
<td>59 (7.0)</td>
<td>60 (7.9)</td>
<td>.57</td>
</tr>
<tr>
<td>Mechanical ventilation, n (%)</td>
<td>590 (70.3)</td>
<td>554 (72.8)</td>
<td>.29</td>
</tr>
<tr>
<td>Length of time (days), mean (SD)</td>
<td>6.1 (9.4)</td>
<td>6.9 (10.3)</td>
<td>.17</td>
</tr>
<tr>
<td>Catecholamine, n (%)</td>
<td>327 (39)</td>
<td>326 (42.8)</td>
<td>.13</td>
</tr>
<tr>
<td>Length of time (days), mean (SD)</td>
<td>3.4 (5.3)</td>
<td>3.7 (7.0)</td>
<td>.41</td>
</tr>
<tr>
<td>Dialysis therapy, n (%)</td>
<td>92 (10.9)</td>
<td>98 (12.9)</td>
<td>.25</td>
</tr>
<tr>
<td>Length of time (days), mean (SD)</td>
<td>7.2 (9.4)</td>
<td>7.6 (11.3)</td>
<td>.79</td>
</tr>
<tr>
<td>All 3 life supports (ventilation, vasopressor, and dialysis), n (%)</td>
<td>56 (6.7)</td>
<td>67 (8.8)</td>
<td>.11</td>
</tr>
</tbody>
</table>

aICCA: Intellispace Critical Care and Anesthesia.

Indirects Markers of Medical Prescriptions

Concerning the markers of medical prescriptions, we noted a significant increase in the volume of prescribed biological analyses, with an increase in the total cost of biological analyses, with a significant increase of hematological and biochemical blood tests. The monthly consumption and cost of biological analyses are summarized in Table 3.
There was no significant difference in the number of chest x-rays prescribed during the 2 study periods (0.55 [SD 0.39] chest x-rays per day of hospitalization per patient before the introduction of ICCA and 0.51 [SD 0.37] chest x-rays per day of hospitalization per patient after the introduction of ICCA; \( P = .05 \)).

There was a decrease in the cost of antibiotic prescription per patient after the introduction of ICCA (€149.50 [$164 USD] per patient before computerization vs €105.40 [$116 USD] per patient after computerization), but the quantity of delivery antibiotics was globally stable in the 2 periods. In addition, in view of the change of markets during the 2 study periods (eg, with a wider use of generics), we decided to repeat the comparison while keeping the cost of antibiotics constant between the 2 periods. We then found a cost per patient of €149.50 ($164 USD) per patient before computerization vs €177.20 ($194 USD) per patient after computerization.

**Table 3.** Monthly consumption and cost of biological analyses.

<table>
<thead>
<tr>
<th>Consumption of biological test</th>
<th>Before ICCA(^a), median (IQR(^b))</th>
<th>After ICCA, median (IQR)</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total volume of biological analyses (number of acts)</td>
<td>5416 (5192-5956)</td>
<td>6374 (6013-6986)</td>
<td>.002(^c)</td>
</tr>
<tr>
<td>Total cost of biological analyses (€)</td>
<td>28,503 (25,531-29,270)</td>
<td>32,530 (30,222-35,973)</td>
<td>.01(^c)</td>
</tr>
<tr>
<td>Volume of hematological blood test (number of acts)</td>
<td>1032 (955-1094)</td>
<td>1182 (1012-1215)</td>
<td>.04(^c)</td>
</tr>
<tr>
<td>Volume biochemical blood test (number of acts)</td>
<td>1628 (1383-1652)</td>
<td>1899 (1675-2062)</td>
<td>.02(^c)</td>
</tr>
<tr>
<td>Volume of procalcitonin and troponin (number of acts)</td>
<td>85 (80-104)</td>
<td>94 (63-123)</td>
<td>.70</td>
</tr>
<tr>
<td>Volume of antibiotic dosage (number of acts)</td>
<td>79 (71-87)</td>
<td>81 (59-96)</td>
<td>.90</td>
</tr>
</tbody>
</table>

\(^a\)ICCA: Intellispace Critical Care and Anesthesia.
\(^b\)IQR: interquartile range.
\(^c\)Significant results.

### Discussion

**Principal Findings**

The implementation of ICIS with ICCA in June 2016 was not associated with a significant modification in the length of stay in our surgical ICU for critically ill patients. These results are the same as those of 2 studies, which reported that after implementation of ICIS in their ICU, the length of stay and morbidity and mortality of hospitalized patients were not affected by the change in prescribing patterns [16,17]. The lack of impact of computerization contradicts the data of Levesque et al who found a reduction in the length of stay of around 2 days after implementation of ICIS, whereas Lyons et al showed an increase in the length of stay in ICU after computerization [14,15]. However, the comparison of these studies remains difficult because each one tested a different software and there is no evidence that these systems are equivalent. To our knowledge, there is no study comparing different types of ICU ICIS software. Given evident logistical constraints, it seems difficult to envisage a randomized study comparing several ICIS software in the same ICU.

We showed that the postcomputerization population was sicker than the precomputerization population, via a significant increase in the severity score SAPSII between the 2 populations. If this difference is statistically significant, we did not consider it clinically relevant. Indeed, the difference between the 2 groups was only 2 points in the severity score, with relatively high scores. The difference was, therefore, small compared with the high average severity scores of our cohorts. One hypothesis is that this may be due to an overall increase in the age of the ICU populations (not demonstrated in our study).

Regarding secondary parameters, very few studies are available. Several studies have evaluated the impact of computerization on the mortality of critically ill patients, particularly for pediatric patients, but they did not show any effect [16-20]. Only a study by Lyons et al showed a reduction in mortality after computerization (3 deaths avoided per 1000 hospitalizations; \( P < .001 \)), which was limited to medical and surgical wards. On the other hand, in this study, the implementation of ICIS was associated with an increase in mortality in ICU [15]. Many studies have focused on demonstrating the safety of computerized surveillance and prescription systems [6-8]. They all showed a reduction in the rate of medication delivery errors and an increase in the rate of errors identified before they were committed [8-10]. However, there is no study looking for a qualitative change of medical prescriptions after the implementation of an ICIS in ICU. We did not analyze every individual medical prescription, and therefore, we did not have direct data concerning prescription errors or practices (eg, forgetting to repeat treatment when prescribing manually). Prescriptions were analyzed indirectly via the quantity of treatments or additional exams consumed. We selected 3 quality markers of medical prescriptions: the consumption of antibiotics (whose duration is fixed and unchanged regardless of the mode of prescription), the consumption of biological analyses (estimated by total cost), and the average number of chest x-rays during the stay. In our study, a reduction in the average cost of antibiotic therapy per patient after implementation of ICCA was observed, but there was no clinically significant difference in the amount of antibiotics delivered. The cost difference was due to a change in market with a large use of generic molecules (as the replacement between the 2 periods of the delivery of Zvyxoid by generic Linzolid) and a diversification of the drugs used (such as the use of fourth-generation cephalosporins and

http://periop.jmir.org/2019/2/e14501/
The implementation of ICCA was associated with a significant increase in the total cost and prescription of biological tests after computerization, whereas the service’s recommendations concerning the prescription of additional tests remained the same over these periods. We observed that the increase in the cost of biological tests did not correspond to the overprescription of some expensive biological tests (eg, troponin or procalcitonin or antibiotics dosages) but corresponded to an increase of biochemical and hematological blood tests prescribed. This over-prescription could have been facilitated by the ease of prescription through a drop-down menu (leading to the prescription of nonrecommended tests). Collin et al who analyzed the impact of the implementation of an information system for medical prescription in 4 general hospitals of the National Health Service in Great Britain observed similar results: the prescription of additional tests was almost quadrupled in patients hospitalized in conventional departments after computerization [21]. This increase of the volume of prescription of additional tests may be due to a change in practices. Indeed, we have the impression that younger generations of physicians used to have easy access to biological and radiological exams, and they seem more likely to rely on complementary exams rather than physical assessment to diagnose and treat patients. A major side effect of this trend is an increase in the cost of care, which is becoming increasingly central to the overall management of patients.

Surveillance and prescription software have already shown a benefit in terms of the amount of care provided to patients by reducing the time spent consulting and collecting the data needed to manage them [24]. Saarinen et al found that computerization in ICU increased the time spent in patient care (81.1% of working time before computerization vs 86.6% after; P<.05) [11]. Similarly, in 2010, Ballermann et al showed a decrease in the time spent consulting the various care documents after setting up a computerized system in their department [12]. However, the efficiency of ICIS in terms of improving the quality of care remains debated. For example, Koppel et al showed that computerized prescription resulted in an increase of 22 types of medication errors, including double prescriptions without computer alerts, confusion of doses and pharmacy inventories, and a lack of overall vision of patient treatments through fragmented computer windows [25]. From indicators such as average length of stay, mortality, and indirect markers of prescription quality, our study suggests that the global quality of care has not been impacted by the implementation of an ICIS. The question that we are raising of whether computerization has a real impact on the prescription of complementary tests and antibiotic therapy could be specified in a future study detailing the individual quality of medical prescription.

Our study faces several strong limitations. First, only the length of stay of the first stay was analyzed. Therefore, each patient was included only once in the analysis. This is a bias in the comparison of the length of stay between the 2 study periods. However, as readmission rates are comparable between the 2 periods, we can consider that this balances the bias caused by the exclusion of successive stays of the same patient.

Furthermore, the quality of medical prescriptions was analyzed indirectly by collecting the quantity of additional biological and radiological tests prescribed as well as the quantity of antibiotics consumed during the 2 study periods. We did not collect and directly analyze prescriptions; hence, some errors such as duplicates, missed retreatments, or dose errors were not taken into account.

Moreover, this retrospective study did not allow us to gather the opinions of caregivers, nurses, and physicians on the computerization of the medical monitoring and prescription system. It has been shown that medical and paramedical caregivers have a positive experience with computerization, but we do not know what the impact of ICCA on our team was [26,27]. Finally, we excluded from our analysis the 2 months following the computerization of medical prescriptions. This may not be sufficient to judge a truly efficient use of the software for both medical and paramedical staff.

Conclusions
On the basis of the indicators collected, the implementation ICCA in our surgical ICU, despite the significant changes it brings about daily practices, did not seem to impact the length of stay in ICU. The real impact of computerization on length of stay and morbidity and mortality in ICU remains controversial. A more detailed approach of medical prescription should better assess the possible benefits of computerized prescription in the specific population of critically ill patients.

Acknowledgments
The authors would like to thank the central pharmacy team, Dr Oudinot, the Institute of Clinical Biology, and Mr Rizo for their assistance in collecting pharmacy and exams data.

Conflicts of Interest
None declared.
Multimedia Appendix 1
Computerized medical prescriptions in ICCA; review of daily prescriptions, listed by route of administration and alphabetical order; example of a new prescription, with drop-down and preconfigured prescription menu, precision of route of administration, dosage, frequency and duration.

[PDF File (Adobe PDF File), 1MB - periop_v2i2e14501_app1.pdf]

Multimedia Appendix 2
ICCA monitoring sheet, showing the different monitored vital parameters, classified by organs (hemodynamics, ventilation...) and function (drainage, analgesia...). The data are automatically validated every 30 minutes, with the possibility of correcting outliers or manually entering data remaining.

[PDF File (Adobe PDF File), 782KB - periop_v2i2e14501_app2.pdf]

References


**Abbreviations**

ICCA: Intellispace Critical Care and Anesthesia  
ICIS: intensive care information system  
ICU: intensive care unit  
SAPSII: Simplified Acute Physiology Score II
Listening to the HysterSisters: A Retrospective Keyword Frequency Analysis of Conversations About Hysterectomy Recovery

Arpit Dave¹, MD; Johnny Yi¹, MD; Andy Boothe², MSSTC; Helene Brashear², PhD; Jeffrey Byrne², BA; Yash Gad², PhD

¹Mayo Clinic Arizona, Department of Gynecology, Phoenix, AZ, United States
²W2O Group, Austin, TX, United States

Abstract

Background: In the postoperative period, individual patient experiences vary widely and are based on a diverse set of input variables influenced by all stakeholders in and throughout the surgical process. Although clinical research has primarily focused on clinical and administrative datasets to characterize the postoperative recovery experience, there is increasing interest in patient-reported outcome measures (PROMs). The growth of online communities in which patients themselves participate provides a venue to study PROMs directly. One such forum-based community is HysterSisters, dedicated to helping individuals through the experience of hysterectomy, a major surgery which removes the uterus. The surgery can be performed by a variety of methods such as minimally invasive approaches or the traditional abdominal approach using a larger incision. The community offers support for “medical and emotional issues [...] from diagnosis, to treatment, to recovery.” Users can specify when and what type of hysterectomy they underwent. They can discuss their shared experience of hysterectomy and provide, among other interactions, feedback, reassurance, sympathy, or advice, thus providing a unique view into conversations surrounding the hysterectomy experience.

Objective: We aimed to characterize conversations about hysterectomy recovery as experienced by users of the HysterSisters online community.

Methods: A retrospective keyword frequency analysis of the HysterSisters Hysterectomy Recovery forum was performed.

Results: Within the Hysterectomy Recovery forum, 33,311 unique users declared their hysterectomy date and type and posted during the first 12 weeks postsurgery. A taxonomy of 8 primary symptom groups was created using a seed list of keywords generated from a term frequency analysis of these threads. Pain and bleeding were the two most mentioned symptom groups and account for almost half of all symptom mentions (19,965/40,127). For symptoms categories such as pain and hormones and emotions, there was no difference in the proportion of users mentioning related keywords, regardless of the type of hysterectomy, whereas bleeding-related or intimacy-related keywords were mentioned more frequently by users undergoing certain minimally invasive approaches when compared with those undergoing abdominal hysterectomy. Temporal patterns in symptom mentions were noted as well. The majority of all posting activity occurred in the first 3 weeks. Across all keyword groups, individuals reporting minimally invasive procedures ceased forum use of these keywords significantly earlier than those reporting abdominal hysterectomy. Peaks in conversation volume surrounding particular symptom categories were also identified at 1, 3, and 6 weeks postoperatively.

Conclusions: The HysterSisters Hysterectomy Recovery forum and other such forums centered on users’ health care experience can provide novel actionable insights that can improve patient-centered care during the postoperative period. This study adds another dimension to the utility of social media analytics by demonstrating that measurement of post volumes and distribution of symptom mentions over time reveal key opportunities for beneficial symptom-specific patient engagement.
Introduction

In the postoperative period, individual patient experiences vary widely and are based on a diverse set of input variables influenced by all stakeholders in and throughout the surgical process [1-3]. Postoperative recovery has been defined as “a dynamic process in an endeavor to continue with everyday life” wherein “individuals strive and struggle to gain independence and return to everyday life” [4]. A total of 4 dimensions of the postoperative recovery process have been described: physiological, psychological, social, and habitual [1]. Although clinical research has focused heavily on the physiological aspect through analysis of clinical and administrative datasets, there is increasing interest in patient-reported outcome measures (PROM), where patients report directly on their own recovery process or experience [5-8]. Standardized validated instruments such as the Quality of Recovery-40 have been developed to elicit this direct information [9]. These PROMs can then ostensibly be used to improve quality of care.

The growth of online communities in which patients themselves participate provides an alternate venue to study PROMs. The authors leveraged this participation to study PROMs in online discussions about postoperative recovery initiated by people who reported undergoing hysterectomy. Hysterectomy is a surgery to remove the uterus and is done for a variety of reasons including leiomyoma (benign smooth muscle tumor), abnormal uterine bleeding, or gynecologic malignancies and can be performed through a variety of surgical approaches [10-13].

The site of the discussions was HysterSisters, an online community dedicated to “issues surrounding the hysterectomy experience […] from diagnosis, to treatment, to recovery” [14]. An active forum dedicated to hysterectomy recovery contains posts from users dating back over 10 years and includes detailed information including the date and type of hysterectomy and conversations about their actual experience. A survey of the HysterSisters community showed that top motivations for posting were obtaining information (87%), experience sharing (76%), and offering advice or information (70%) [15].

Previous work on HysterSisters administered a Likert-based satisfaction survey to a self-selected population of forum participants [16]. The survey covered a variety of general postoperative recovery domains including overall hysterectomy results, time to return to normal activity, pain and discomfort, and others. Here, we build on this work by performing a linguistic analysis of the posts from the population of site users who participated in the posthysterectomy recovery forum.

There are numerous techniques with which to approach the content analysis of online forum posts [17,18]. The choice of tool is dependent on the information sought. Tools such as Linguistic Inquiry and Word Count use preset dictionaries to assist in tasks such as sentiment analysis, whereas use of a topic modeling strategy such as latent Dirichlet allocation or a simpler term frequency analysis can identify topics of conversation within a bag-of-words corpus of forum text [19,20].

This study attempted to characterize more completely the recovery experience of individuals in the HysterSisters community by studying the subject headings of users’ own publicly available posts to listen to conversations outside the provider’s office. Therefore, the objective was to identify conversation patterns in forum data to provide actionable insights into the surgical recovery experience.

Methods

Study Design

In accordance with the Code of Federal Regulations 45 CFR 46, the Mayo Clinic Institutional Review Board (IRB) deemed this study does not require IRB review on the basis that publicly accessible contributions to the HysterSisters forums do not constitute private behavior. No registration or login was required to read posts.

We performed a retrospective mixed methods analysis of forum posts that leveraged both structured and free-text user inputs. Qualitative assessments included development of a symptom taxonomy of forum topic keywords that encompass the 4 dimensions of recovery as described previously. Although each keyword may not represent a medical or pathologic symptom and should not be equated with a patient complaint, we use the word in its broader context as an indicator of an element of the general recovery experience. Quantitative investigations included keyword frequency analyses and survival analyses of these keywords, each of which is described below.

Data Collection

The HysterSisters website provided users with the option to enter structured data including type of hysterectomy, ovarian status, and exact date of procedure [8]. The forum’s taxonomy of hysterectomy types was categorized for analysis into treatment groups by surgical approach (Textbox 1). As there are many approaches to hysterectomy, the question of which approach is superior is of interest, as are reasons a surgeon or patient might choose one approach over another. The American College of Obstetrics and Gynecology considers both the vaginal hysterectomy (VH) and laparoscopic hysterectomy (LH) to be minimally invasive, with well-described benefits over the AH approach, and a recent Cochrane systematic review also demonstrated favorable clinical outcomes for both VH and LH when compared with AH [21,22]. We chose a similar organizational scheme to study the question of hysterectomy recovery experience from the forum user’s own words.
Although posts in the board were public, data collection of whether ovaries were removed at the time of hysterectomy. We did not exclude any users based on reporting therefore chose to look out 12 weeks from the reported date of hysterectomy. We sought to determine if users continued to engage beyond this time frame, given that (1) anecdotal evidence suggests that patients experience subclinical complications of surgery (eg, bloating, discomfort, and hormonal symptoms) beyond 6 weeks, and (2) a major complication of hysterectomy (dehiscence of the vaginal cuff) has been demonstrated to have a median time to occurrence as long as 11 weeks [24].

We have shown that median time to return to work may be upward beyond 6 weeks, and (2) a major complication of hysterectomy is thought to last 6 weeks, after which patients are typically discharged to routine annual follow-up. However, the research by Vonk Noordegraaf et al has shown that median time to return to work may be upward of 8 weeks [23]. We sought to determine if users continued to engage beyond this time frame, given that (1) anecdotal evidence suggests that patients experience subclinical complications of surgery (eg, bloating, discomfort, and hormonal symptoms) beyond 6 weeks, and (2) a major complication of hysterectomy (dehiscence of the vaginal cuff) has been demonstrated to have a median time to occurrence as long as 11 weeks [24]. We therefore chose to look out 12 weeks from the reported date of hysterectomy. We did not exclude any users based on reporting of whether ovaries were removed at the time of hysterectomy. Although posts in the board were public, data collection was preserved anonymity by processing forum text and metadata without the username.

### Symptom Keyword Frequency Analysis

Term frequency analysis is an analytical technique that characterizes the differences between 2 text corpora by comparing the relative frequencies at which n-grams appear in each corpus [25,26]. N-grams are a contiguous sequence of n words, where n is an integer. This n-gram analysis used a log-likelihood approach to term frequency analysis because it offers a test for significance [27]. To gain understanding of how users in different treatment groups discussed the same symptom differently, 1 corpus each was constructed from the subjects of the posts mentioning the given symptom for the 2 treatment groups to compare, and then term frequency analysis was applied to compare the 2 corpora. For this analysis, the abdominal treatment group subject headers were used as the base corpus, against which each minimally invasive surgery group (laparoscopic and vaginal) was individually compared. The rationale for this comparison emerges from the consensus in clinical gynecology that the minimally invasive treatments are preferable to traditional AH because of less surgical risk and faster recovery, as we sought to identify if these patterns emerged in online conversations as well [21,22].

This analysis yielded a list of n-grams used more by each treatment group when mentioning each symptom versus the abdominal treatment group; these lists prioritized and motivated any manual, qualitative review of matching posts. All comparisons used equality to the abdominal treatment group’s mention frequency as the null hypothesis to allow comparison of the set of minimally invasive treatments with the traditional treatment.

### Symptom Taxonomy

To translate agnostic forum text into clinically meaningful information, a symptom keyword taxonomy was developed using a subjective, iterative, collaborative process between the medical and computational researchers. The initial keywords were identified using an n-gram (n=1, 2, and 3) frequency analysis of all posts in the Hysterectomy Recovery board, which

---

**Textbox 1.** HysterSisters website taxonomy for hysterectomy (available hysterectomy types categorized into treatment groups by surgical approach).

<table>
<thead>
<tr>
<th>Abdominal</th>
<th>Vaginal (minimally invasive surgery)</th>
<th>Laparoscopic (minimally invasive surgery)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Total abdominal (TAH)</td>
<td>1. Total vaginal (TVH)</td>
<td>1. Total laparoscopic (TLH)</td>
</tr>
<tr>
<td>2. Supracervical abdominal (SAH)</td>
<td>2. Laparoscopic-assisted vaginal (LAVH)</td>
<td>2. Laparoscopic supracervical (LSH)</td>
</tr>
<tr>
<td>3. Either, not specified (TAH/SAH)</td>
<td>3. da Vinci robotic laparoscopic (DVH)</td>
<td>3. da Vinci robotic laparoscopic (DVH)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Single-incision laparoscopic (SILS or laparoendoscopic single-site surgery [LESS])</td>
</tr>
</tbody>
</table>
revealed a seed list of commonly reported words used to describe symptoms. The list was then expanded by alternately searching by keywords and inspecting text that was both included and excluded by search queries. Keyword searches were used to pull subject headers containing the keyword and then new keywords were added to the list after manually examining the conversations beneath that header. This process continued until the keyword searches maintained a consistent conversation indexing. Individual keywords were chosen to prioritize specificity over sensitivity; sensitivity was maximized by including many keywords.

The final keyword symptom groups emerged as pain, sleep and fatigue, hormones and emotions, digestion, swelling, bleeding, urination, intimacy, odd sensations, drugs, fever and infection, and family. The set of all words for the keyword symptom groups included technically accurate terms (gastritis), proper English words (eg, itches and burns), slangs (eg, weepies and swellybelly), and common typographical and spelling errors (eg, achey and vomiting). Multimedia Appendix 1 shows the keyword symptom groups and all included words, organized alphabetically, and Multimedia Appendix 2 shows the same, sorted by number of mentions in descending order.

**Symptom Keyword Mention Frequency Analysis**

The symptom mention frequency analysis compared the number of users who mentioned each symptom broken out by treatment group. The subject headers were tagged for symptom mentions by searching tokenized post subjects for the corresponding symptom keywords from the taxonomy. Responses were aggregated by user to compile the list of symptoms each user mentioned during the 12-week postoperative recovery period. A chi-square test for homogeneity was used to compare the mention frequencies of each symptom among individuals in each hysterectomy group with the abdominal group.

**Symptom Keyword Mention Survival Analysis**

To compare whether individuals undergoing minimally invasive treatments stopped discussion of symptoms on the forum earlier than their counterparts undergoing AH, the same set of tagged subject headers was grouped by user and sorted chronologically to determine each user’s latest mention of each symptom. Those who did not mention a symptom were excluded from the survival analysis of that symptom only; in the event that more than 1 symptom category was mentioned by a user, each symptom category was considered separately. A log-rank test was used to compare the final mentions of a symptom among each treatment group. All comparisons used equality to the abdominal treatment group’s corresponding symptom survival curve as the null hypothesis. Mean interquartile difference between survival curve pairs were calculated to quantify which group ceased to mention symptom keywords earlier.

**Software**

All data processing and analysis were done using free, open-source libraries written in Python (Python Software Foundation). Data processing and aggregation were performed using the pandas library, text processing with the nltk library, and statistical analysis using the scipy library.

**Results**

**Summary Statistics**

There were 33,311 unique users in the Hysterectomy Recovery forum, making at least one mention of a symptom in the taxonomy. Among these contributors, the procedure distribution is as follows: abdominal=13,306/33,311 (39.94%), vaginal=10,589/33,311 (31.79%), and laparoscopic=9,416/33,311 (28.27%). Among users who provided ovary status data, there were more who kept at least one ovary (18,645/33,311, 55.97% of all users) than who had both removed (12,313/33,311, 36.96%); some did not specify their ovary status (2,353/33,311, 7.06%). Ovary status by treatment group is shown in Figure 1.

**Conversation Volume**

Site users with completed profiles created a total of 80,704 top-level posts during the first 12 weeks of their respective recoveries. The subjects of 42.43% (34,242/80,704) of these posts mentioned at least one symptom as defined by the symptom taxonomy; the remaining 57.57% (46,462/80,704) mentioned none.

Posting behavior was heavily skewed, with most posts (42,910/80,704, 53.17%) happening within the first 3 weeks of the 12-week recovery period being studied. Figure 2 shows a histogram of posts, segmented by days postoperation and stratified by symptom mention count. The median post was made during day 19 ($\mu$=day 23.80 and $\sigma$=18.56 days). The median posts per contributor was 1 post ($\mu$=2.42 posts and $\sigma$=3.10 posts).

**Symptom Mention Volume**

Multimedia Appendix 3 shows a bar chart of users who mention each symptom. The top 3 symptoms by volume of mentions are pain, bleeding, and hormones and emotions for both the aggregate conversation and each of the conversations by procedure.

**Sorting by Relevance**

The 34,242 subject headers, which mention at least one symptom contain 40,127 total symptom mentions. The symptoms mentioned the most were pain (12,474/34,244, 36.43% of subject headers that mention at least 1 symptom) and bleeding (7,491/34,242, 21.88%); together, these symptoms account for half of all symptom mentions. Relevant posting behavior follows overall posting volume very closely. Proportional symptom mentions per unit time remain generally flat throughout the first 12-weeks of recovery with few important exceptions (Figure 3).
Figure 1. Ovary status by surgical approach. Total number of procedures reported by HysterSisters patients mentioning at least one symptom included in the taxonomy, stratified by surgical approach. Each procedure is broken down by ovary status. Unknown indicates patients did not provide ovary status data.

Figure 2. Post volume by days postoperative. Bars indicate total number of posts created by HysterSisters patients, grouped by the number of whole days postoperation the post was created. Bars are broken down by number of symptoms each post subject mentions according to the symptom taxonomy.
User Symptom Keyword Mention Frequency Analysis

There is a significant difference between the number of users who mention a symptom at any point during recovery for a given treatment group versus the abdominal group for some symptoms. For example, users attesting to VH mentioned urinary and intimacy keywords proportionally more. Users having had an LH mentioned bleeding proportionally more. Both mentioned swelling and sleep and fatigue-related keywords less. There were no differences in the frequency of mentions of pain and hormone and emotion keywords. Table 1 shows the absolute percentage difference in mentions for each treatment group and symptom permutation versus the abdominal surgical group.

Table 1. Absolute percentage difference for mentions of a given symptom by procedure compared with the abdominal group.

<table>
<thead>
<tr>
<th>Symptom keyword</th>
<th>Laparoscopic (%)</th>
<th>Vaginal (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>-0.05(^a)</td>
<td>0.38(^b)</td>
</tr>
<tr>
<td>Drugs</td>
<td>-0.81(^a)</td>
<td>-0.60(^a)</td>
</tr>
<tr>
<td>Urinary</td>
<td>-0.50(^a)</td>
<td>2.29(^{b,c})</td>
</tr>
<tr>
<td>Hormones and emotions</td>
<td>-0.49(^a)</td>
<td>0.28(^b)</td>
</tr>
<tr>
<td>Intimacy</td>
<td>-0.25(^a)</td>
<td>2.38(^{b,c})</td>
</tr>
<tr>
<td>Sleep and Fatigue</td>
<td>-2.22(^{a,d})</td>
<td>-1.69(^{a,e})</td>
</tr>
<tr>
<td>Swelling</td>
<td>-2.94(^{a,d})</td>
<td>-3.04(^{a,c})</td>
</tr>
<tr>
<td>Pain</td>
<td>0.21(^b)</td>
<td>0.98(^b)</td>
</tr>
<tr>
<td>Fever and infection</td>
<td>-0.42(^a)</td>
<td>0.39(^b)</td>
</tr>
<tr>
<td>Digestion</td>
<td>1.53(^b)</td>
<td>0.41(^b)</td>
</tr>
<tr>
<td>Odd sensations</td>
<td>-0.43(^a)</td>
<td>-0.96(^a)</td>
</tr>
<tr>
<td>Bleeding</td>
<td>3.03(^{a,d})</td>
<td>-1.46(^b)</td>
</tr>
</tbody>
</table>

\(^a\)Values indicate the abdominal cohort mentions the symptom more.
\(^b\)Values indicate the abdominal cohort mentions the symptoms less.
\(^c\)\(P<.001\).
\(^d\)\(P<.01\).
\(^e\)\(P<.05\).
User Symptom Keyword Mention Survival Analysis

Users in the minimally invasive treatment groups ceased to mention nearly all of the symptoms being studied significantly earlier versus the abdominal group, even in cases where proportionally more users mentioned the symptom. For example, users in the LH group ceased to mention bleeding at a mean interquartile difference of 1.66 days sooner ($P=.003$) than in AH group. In the VH group, users ceased to mention pain keywords at a mean interquartile difference of 4.00 days sooner ($P<.001$) than in the AH group. Table 2 lists the mean interquartile differences in days between the timing of cessation of each symptom mention by treatment group compared with the AH cohort.

Table 2. User symptom keyword survival mean interquartile differences versus abdominal group, in days.

<table>
<thead>
<tr>
<th>Symptom keyword</th>
<th>Laparoscopic</th>
<th>Vaginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>$-9.3^{a,b}$</td>
<td>$-5.0^{a,c}$</td>
</tr>
<tr>
<td>Drugs</td>
<td>$-3.6^{a,c}$</td>
<td>$-1.3^{a}$</td>
</tr>
<tr>
<td>Urinary</td>
<td>$-7.6^{a,b}$</td>
<td>$-4.6^{a,b}$</td>
</tr>
<tr>
<td>Hormones and emotions</td>
<td>$-5.6^{a,b}$</td>
<td>$-5.6^{a,b}$</td>
</tr>
<tr>
<td>Intimacy</td>
<td>$-0.3^{a}$</td>
<td>$1^{d}$</td>
</tr>
<tr>
<td>Sleep and fatigue</td>
<td>$-6.3^{a,b}$</td>
<td>$-3.6^{a,c}$</td>
</tr>
<tr>
<td>Swelling</td>
<td>$-7.0^{a,b}$</td>
<td>$-4.0^{a,c}$</td>
</tr>
<tr>
<td>Pain</td>
<td>$-7.0^{a,b}$</td>
<td>$-4.0^{a,b}$</td>
</tr>
<tr>
<td>Fever and infection</td>
<td>$-7.3^{a,b}$</td>
<td>$-6.0^{a,b}$</td>
</tr>
<tr>
<td>Digestion</td>
<td>$-4.0^{a,b}$</td>
<td>$-2.3^{a,b}$</td>
</tr>
<tr>
<td>Odd sensations</td>
<td>$-4.3^{a,b}$</td>
<td>$-2.6^{a,c}$</td>
</tr>
<tr>
<td>Bleeding</td>
<td>$-1.6^{a,c}$</td>
<td>$-2.0^{a,b}$</td>
</tr>
</tbody>
</table>

$^{a}$Values indicate an earlier cessation of mentions.  
$^{b}$P<.001.  
$^{c}$P<.01.  
$^{d}$Values indicate later cessation of mentions.

Discussion

Results Analysis

The HysterSisters forum dataset provides an opportunity not only to broadly sample patient online conversations regarding hysterectomy recovery, benchmarked by date, and type of procedure but also, more broadly, a method by which online conversations can be used to inform perioperative care for similar communities surrounding different clinical experiences.

This analysis provides rich insight into the hysterectomy recovery experience. First, the temporal dynamics of individual engagement on the forum are quite varied. Individuals seek engagement most heavily in the first 3 weeks after hysterectomy. However, there are also specific windows during the recovery in which engagement is desired. For hysterectomy, these peaks in conversation volume occur at 1, 3, and 6 weeks postoperative. In addition, the topic of interest changes as well. During the first week, digestion issues are of considerable concern, but at 1 week and beyond, a relative increase in the percentage of bleeding mentions suggests bleeding as a potential focus of assessment, reassurance, or counseling for patients. Conversations related to intimacy arise starting at 3 weeks, spiking at 6 weeks postoperatively, coinciding with providers’ approval to return to sexual activity; however, this spike may also suggest the presence of a persistent information gap patients seek to fill. That the distribution of keyword mentions remains otherwise constant throughout the 12-week recovery indicates users do continue to desire engagement on all these topics throughout and beyond the standard recovery period. Investigators pursuing similar avenues of research should consider such dynamism when analyzing posting behavior.

Second, procedural variations should be accounted for as they may impact the clinical applications of the research. In our case, the type of hysterectomy was captured as structured data. Individuals undergoing VH make proportionally more mentions of urinary symptoms. This difference in mentions may be because of more women in this cohort undergoing concomitant prolapse or incontinence surgeries, which this analysis does not explore. Clearly, however, patients seek engagement here, and addressing urinary function can help maximize patient satisfaction with their recovery experience. We also noted topics about bleeding occurred more frequently from individuals undergoing the various laparoscopic hysterectomies. This difference in conversation frequency between approaches was in contrast to clinical data presented in the Cochrane review on surgical approach to hysterectomy where no evidence of a difference in the number of individuals with substantial bleeding between laparoscopic and AH groups was seen [22]. Therefore,
we undertook a manual investigation into the posts. A 3-gram analysis of the laparoscopic cohort noted at weeks post is mentioned significantly more ($P<.001$), and examples include “Vaginal discharge at 10 weeks post hysterectomy,” “New slight spotting at 10 weeks post hysterectomy,” and “Spotting and slight pain at 10 weeks post hysterectomy.” These examples suggest patients may experience resumption of bleeding after a perceived recovery. Although dissolution of delayed absorbable suture is a ready explanation in this instance, the example demonstrates how this type of research reveals opportunities for anticipatory guidance.

Notably, although there are no significant differences in the frequency of mentions of pain-related symptoms, the survival analysis shows the last mention of pain occurring about 7 days earlier for the laparoscopic cohort with about half of these users ceasing to mention pain by postoperative day 20. It is tempting to interpret this finding as half of the users stop experiencing pain at 3 weeks, but this may not be the case. More appropriately, as users’ experience begins to match their expectations for pain or any particular symptom, the need to engage socially may diminish. As noted above, reengagement can occur when expectation-experience mismatching occurs.

Nevertheless, with the exception of intimacy-related keywords, cessation of symptom mentions occurs earlier in the vaginal and laparoscopic cohorts across all symptom groups. Therefore, although our analysis is not intended to deliver concrete recommendations as to the route of hysterectomy, the findings do parallel those in clinical gynecology literature where return to normal activities was found to occur earlier after vaginal and laparoscopic hysterectomies versus the abdominal approach [22].

Taken together, these findings can help guide clinical postoperative care. For example, the interquartile difference in days for cessation of bleeding mentions is only 2.0 days for the vaginal compared with AH groups; therefore, practically, bleeding should be discussed regardless of hysterectomy type and remains a concern throughout the recovery process. In addition, gastrointestinal and genitourinary symptoms should receive focus early (at discharge), whereas providers should be sure to address intimacy issues at the final postoperative visit and reassure patients of their ongoing availability for care as the patients’ needs may continue past the typical 6-week clinical recovery period. We present a simple reminder chart to alert providers to review these critical topics (Textbox 2).

**Conversational Perspectives**

Although internet search and online forum usage are rising among patients, patients still overwhelmingly turn to their doctor for medical expertise. In 1 study, 91% of patients sought their doctor for medical diagnosis. However, when asked about practical advice for coping with day-to-day health situations, patients were more divided with 43% choosing their doctor and 46% choosing the group including fellow patients, friends, and family [28].

This analysis has important a priori limitations. First, symptoms keywords were sorted into taxonomies by a subjective iterative process because the authors must ultimately assign any keyword to a symptom group. The taxonomy may be incomplete because of sparseness of typographical error, unanticipated slang, or proper medical terminology reported unassociated with a symptom (eg, catheter). Our taxonomy is therefore included as an appendix.

Second, this analysis should not be seen as equating mentions of symptoms with patient complaint. Deeper analysis could begin to explore subject headers, post content, and conversational patterns related to motivation.

Finally, almost 60% of posts in the Hysterectomy Recovery board did not fall into our taxonomy. Excluded subject headers include a variety of content, including nonsymptom issues (eg, return to activity, comorbid condition issues, or nonmedical topics such as makeup), progress updates, or chatter.

**Future Directions**

This study presents a first look into text analytics to explore the patient experience in gynecologic care and how such research might be conducted in other fields. Focus is limited, however, to the subject headers of posts by site users. Future research analyzing the post body itself and other conversational elements and patterns, we can begin to ascertain the underlying motivation for posting.

**Textbox 2.** Clinician’s quick guide to hysterectomy postoperative counseling.

| At discharge |
| Discuss: gastrointestinal and genitourinary function; typical delay in return of normal bowel function; symptoms of urinary retention versus return of normal voiding; proper pain medication use; bleeding expectations |
| At 1 week |
| Discuss: reassure that intermittent light bleeding is normal if present; ensure adequate return of bowel and bladder function; readdress pain control |
| At 3 weeks |
| Discuss: general check-in with patient; identify individual issues; address hormones, emotional changes, and coping strategies |
| At 6 weeks |
| Discuss: return to normal activity; return to sexual activity; address hormonal changes if persistent or indicated; review possibility of light bleeding 10 weeks postoperative |

Our results demonstrate a timeline of posts with shifting conversational volume in specific areas, with the majority of posts occurring in the first 2 weeks. Future research can integrate this information into existing personalized electronic health
programs for recovery from gynecologic surgery to deliver just-in-time information to patients [29]. Push notifications have been used clinically for such active engagement [28,30,31]. We hope that our results provide insight to both the gynecologic surgeon as to what their patients are discussing after hysterectomy and the data scientist using this information to better analyze similar text-based data sets in other fields.

Acknowledgments
No financial support was solicited or used for the conduct of this research. Coauthors from W2O Group provided computational resources and statistical analysis services at no cost to the authors and participated in the preparation of the Methods and Results sections.

Conflicts of Interest
Authors AB, JB, and HB are employees of W2O Group. YG was an employee of W2O Group at the time of data collection and initial submission. W2O Group lists Intuitive Surgical (Sunnyvale, CA) as a client. No funding from Intuitive Surgical was obtained for this study. W2O Group subsidiaries offer commercial data analytics and marketing services in numerous economic sectors, including health care. AD was a surgical fellow at Mayo Clinic at the time of data collection and initial submission.

Multimedia Appendix 1
Symptom keyword taxonomy, alphabetical.

Multimedia Appendix 2
Symptom keyword taxonomy, by mentions.

Multimedia Appendix 3
Patients who mention a symptom during the recovery period. Each patient is counted at most once per symptom, even if the symptom was mentioned multiple times.

References


Abbreviations

AH: abdominal hysterectomy