Contents

Original Papers

Discharge Instruction Reminders Via Text Messages After Benign Gynecologic Surgery: Quasi-Experimental Feasibility Study (e22681)
Jocelyn Sajnani, Kimberly Swan, Sharon Wolff, Kelsi Drummond ................................................................. 3

Sustaining a Multidisciplinary, Single-Institution, Postoperative Mobilization Clinical Practice Improvement Program Following Hepatopancreatobiliary Surgery During the COVID-19 Pandemic: Prospective Cohort Study (e30473)
Kai Chan, Bei Wang, Yen Tan, Jaclyn Chow, Ee Ong, Sameer Junnarkar, Jee Low, Cheong Huey, Vishal Shelat. ............................................................ 22

Discussion of Weight Loss Surgery in Instagram Posts: Successive Sampling Study (e29390)
Zoe Meleo-Erwin, Corey Basch, Joseph Fera, Bonnie Smith ............................................................................. 33

Patient Engagement in the Design of a Mobile Health App That Supports Enhanced Recovery Protocols for Cardiac Surgery: Development Study (e26597)
Anna Chudyk, Sandra Ragheb, David Kent, Todd Duhamel, Carole Hyra, Mudra Dave, Rakesh Arora, Annette Schultz ............................................................................. 42

Mobile Health Apps That Act as Surgical Preparatory Guides: App Store Search and Quality Evaluation (e27037)
Naga Gadde, Kevin Yap ........................................................................................................................................ 50

The Psychological Experience of Frontline Perioperative Health Care Staff in Responding to COVID-19: Qualitative Study (e27166)
Toni Withiel, Elizabeth Barson, Irene Ng, Reny Segal, Daryl Williams, Roni Krieser, Keat Lee, Paul Mezzavia, Teresa Sindoni, Yinwei Chen, Caroline Fisher ........................................................................................................................................ 59

Management of Acute Appendicitis in Children During COVID-19 and Perspectives of Pediatric Surgeons From South Asia: Survey Study (e26613)
Md Hannan, Mosammat Parveen, Md Hoque, Tanvir Chowdhury, Md Hasan, Alak Nandy .................................................................................... 74

Predicting Prolonged Apnea During Nurse-Administered Procedural Sedation: Machine Learning Study (e29200)
Aaron Conway, Carla Jungquist, Kristina Chang, Navpreet Kamboj, Joanna Sutherland, Sebastian Mafeld, Matteo Parotto ............................................................................. 83
Commentary

Context and Complexity in Telemedicine Evaluation: Work Domain Analysis in a Surgical Setting (e26580)
Hedvig Aminoff, Sebastiaan Meijer

9
Discharge Instruction Reminders Via Text Messages After Benign Gynecologic Surgery: Quasi-Experimental Feasibility Study

Jocelyn Sajnani¹,², MD; Kimberly Swan³, MD; Sharon Wolff⁴, MPH; Kelsi Drummond⁵, MD

¹University of Kansas School of Medicine, University of Kansas Medical Center, Kansas City, KS, United States
²Department of Obstetrics and Gynecology, University of California, Los Angeles, Los Angeles, CA, United States
³Department of Obstetrics and Gynecology, Overland Park Regional Medical Center, Overland Park, KS, United States
⁴Department of Population Health, University of Kansas Medical Center, Kansas City, KS, United States
⁵Department of Obstetrics and Gynecology, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:
Jocelyn Sajnani, MD
Department of Obstetrics and Gynecology
University of California, Los Angeles
10833 Le Conte Ave
CHS 27-139
Los Angeles, CA, 90025
United States
Phone: 1 310 825 9945
Email: jrsajnani@gmail.com

Abstract

Background: With the implementation of enhanced recovery after surgery protocols and same-day hospital discharge, patients are required to take on increasing responsibility for their postoperative care. Various approaches to patient information delivery have been investigated and have demonstrated improvement in patient retention of instructions and patient satisfaction.

Objective: This study aimed to evaluate the feasibility of implementing a postoperative text messaging service in the benign gynecologic population.

Methods: We used a quasi-experimental study design to evaluate patients undergoing outpatient laparoscopic surgery for benign disease with a minimally invasive gynecologist at an academic medical center between October 2017 and March 2018. In addition to routine postoperative instructions, 19 text messages were designed to provide education and support to postoperative gynecologic patients. Patients were contacted by telephone 3 weeks postoperatively and surveyed about their satisfaction and feelings of connectedness during their recovery experience. Demographic and operative information was gathered through chart review. The cost to implement text messages was US $2.85 per patient.

Results: A total of 185 patients were eligible to be included in this study. Of the 100 intended intervention participants, 20 failed to receive text messages, leaving an 80% success in text delivery. No patients opted out of messaging. A total of 28 patients did not participate in the postrecovery survey, leaving 137 patients with outcome data (control, n=75; texting, n=62). Satisfaction, determined by a score ≥9 on a 10-point scale, was 74% (46/62) in the texting group and 63% (47/75) in the control group (P=0.15). Connectedness (score ≥9) was reported by 64% (40/62) in the texting group compared with 44% (33/75) in the control group (P=0.02). Overall, 65% (40/62) of those in the texting group found the texts valuable (score ≥9).

Conclusions: Postoperative text messages increased patients’ perceptions of connection with their health care team and may also increase their satisfaction with their recovery process. Errors in message delivery were identified. Given the increasing emphasis on patient experience and cost effectiveness in health care, an adequately powered future study to determine statistically significant differences in patient experience and resource use would be appropriate.

(JMIR PeriOper Med 2021;4(2):e22681) doi:10.2196/22681

KEYWORDS
communication; hysterectomy; minimally invasive; laparoscopy; postoperative; patient satisfaction
Introduction

With the increase in same-day hospital discharge, patients are required to take on increasing responsibility for their postoperative care. Low health literacy has been consistently associated with poor health [1]. Adequate health literacy is required to follow discharge instructions and has a significant impact on the care of surgical patients [2]. Yet, despite patient education prior to discharge, patients continue to have questions about routine postoperative care when at home [3]. This may represent patient difficulty in retaining medical information, as patients tend to focus more on information related to their diagnosis at the expense of instructions regarding treatment [4]. The lack of information retention and associated poor outcomes underline the importance of continued improvement in patient education regarding discharge care.

Various approaches to patient information delivery have been investigated and have demonstrated improvement in patient retention of instruction. Successful studies have used pictographs [5], multimedia video [6], and other electronic reminders like text messages [7]. Text messaging may be particularly beneficial for the postoperative surgical population who have unique medical needs and require robust education about their postoperative care. A previous feasibility study examining patients undergoing breast reconstruction and anterior cruciate ligament repair found implementing a mobile application for monitoring quality of recovery at home was feasible and acceptable to patients [8].

Previous research can be leveraged and applied to the benign gynecologic postsurgical population through carefully curated text messages that provide education and support during the postoperative recovery period. The intent of this study was to evaluate the feasibility of implementing a postoperative text messaging service.

Methods

Sample and Messaging

This quasi-experimental study included women undergoing benign gynecologic, laparoscopic surgery with a single minimally invasive gynecologist at an academic medical center over a 6-month period. The institutional review board approved this project with quality improvement determination. Patients were eligible for inclusion if they were at least 18 years old and their primary language was English. Those patients undergoing open procedures or laparoscopic procedures converted to open were excluded. For a 3-month period, the control group received routine discharge instructions after their procedures. During the subsequent 3-month period, the intervention group received routine discharge instructions in addition to text messages on days 1 through 8 following their procedure. These messages were implemented using an existing commercially available service for appointment reminder text messaging via TeleVox Solutions. Phone numbers for patients meeting inclusion criteria were pulled directly from patients’ charts. Initial attempt at messaging was made with listed mobile phone numbers. If the number was missing or inaccurate, a second attempt was made using listed home phone numbers, with the understanding that home phone numbers are often mobile numbers. In total, 19 text messages were transmitted to patients via one-way messaging, using their procedure date as an anchoring point. The messages were written to address common postoperative milestones, provide recovery tips, identify situations in which to escalate care, and lastly, to provide encouragement during the recovery process (Figure 1). Patients were universally opted into automated messaging, however, could elect to opt out upon receipt of the initial text message. The cost to implement text messaging was US $0.15 per text, which was the rate negotiated with TeleVox Solutions, with a total cost of US $2.85 per patient for the complete text series. Departmental research funds were used to cover the cost of messaging. Patients in both the control and texting groups were contacted 3 weeks postoperatively for participation in a phone survey. They were asked to rate the following on a scale of 1 to 10: (1) “How satisfied are you with your postoperative care?” (2) “How connected did you feel to your health care team while recovering at home?” (3) “How valuable did you find the text messages while recovering at home?”

The third question was only posed to the texting group. Additionally, demographic, procedural, and postoperative complication data were collected through chart review.
Outcome Measures

As a pilot study, the primary intent was to determine the feasibility of implementing a text messaging protocol and successfully sending messages to patients. The primary endpoint was the percent of intended intervention patients who received all messages without error. Establishing that the intervention did not harm and might benefit patient care was necessary to justify continued evaluation of this intervention. Therefore, important secondary endpoints were patient scores on satisfaction, feelings of connection, and value of text messaging during their postoperative recovery. Satisfaction, connection, and value of text messaging were collected using a 10-point scale. These scores were collapsed into binary variables with scores of 9 and 10 coded as being satisfied, connected, and finding value. Scores of 1 through 8 represented a neutral or negative response. This determination was largely based on clinical importance.

Statistical Analysis

Descriptive statistics were calculated for all outcome and clinical variables. The primary outcome was calculated as the proportion of patients who received all text messages out of those patients who were intended to receive messaging. Chi-squared analyses were used to compare patient characteristics as well as the outcomes of satisfaction and connectedness between the control and intervention groups. To determine if age modified the relationship between texting and the outcomes of satisfaction and connectedness, a Breslow-Day test was conducted. The median age of the cohort (37 years) was used to define younger (<37 years) and older (≥37 years) subgroups.

Results

Text Message Delivery Success Rate

A total of 185 patients were identified to be included in this study. In the texting group, 20 patients did not receive the initial text message due to an error, likely due to the use of a landline phone number. In addition, 10 control and 18 intervention patients did not participate in the postrecovery survey, leaving 137 patients with outcome data (control, n=75; texting, n=62). The primary endpoint was the percent of intended intervention patients who received all messages without error. Of the 100 intended texting group patients, 20 did not receive messages due to error, leaving 80% (80/100) of the intended intervention patients successfully receiving all 19 text messages. Of note, no patients receiving messaging opted out after receiving the initial text, which included the choice to opt out (Figure 2).
Participant Demographic and Clinical Data

The age of study participants ranged from 20 years to 59 years, with a mean age of 38.1 years (SD 8.8). Control and intervention patients were statistically different in terms of age ($P=.04$), with the mean age of the texting cohort being 2.9 years younger than that of the control group. Otherwise, the groups did not differ significantly in terms of other procedural or clinical metrics (Table 1).

### Table 1. Demographic and clinical characteristics of the control and intervention groups.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control group (n=75)</th>
<th>Texting group (n=62)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>39.5 (8.6)</td>
<td>36.4 (8.7)</td>
<td>.04</td>
</tr>
<tr>
<td>Overnight stays, n (%)</td>
<td>21 (28)</td>
<td>11 (18)</td>
<td>.16</td>
</tr>
<tr>
<td>Emergency department visit, n (%)</td>
<td>4 (5)</td>
<td>4 (7)</td>
<td>.78</td>
</tr>
<tr>
<td>Hysterectomy, n (%)</td>
<td>37 (49)</td>
<td>24 (39)</td>
<td>.21</td>
</tr>
</tbody>
</table>

Patient Satisfaction With Postoperative Recovery

Based on completed survey responses, 63% (47/75) of patients receiving standard discharge instructions were satisfied with their postoperative care, while 74% (46/62) of patients receiving additional educational text messaging were satisfied ($P=.15$). Although the texting group patients were found to be younger, age did not modify patient-reported satisfaction scores. Patients <37 years old in the texting group rated their satisfaction similar to those ≥37 years old in the texting group (75%, 47/62 vs 73%, 55/75; $P=.86$). Type of operative procedure, categorized as hysterectomy or uterine-sparing, did play a role in patient satisfaction. For those in the control group with standard written instructions, 76% (28/37) of those patients undergoing hysterectomy were satisfied with their postoperative care, while only 50% (19/38) undergoing uterine-sparing procedures were satisfied ($P=.02$). In contrast, this relationship between type of procedure and satisfaction did not exist in the texting group ($P=.91$).

Patient Connection With the Health Care Team

Survey responses showed that 64% (40/62) of patients receiving educational and supportive text messages felt more connected with their health care team while recovering at home compared with 44% (33/75) of patients receiving only written discharge instructions ($P=.02$). Similar to patient satisfaction, age below or above the median did not alter reported connection scores in the texting group ($P=.34$). The type of procedure did influence the patient perception of connectedness. Patients in the texting group reported similar rates of connectedness regardless of type of procedure performed ($P=.78$). In contrast, those in the control group who underwent hysterectomy felt much more connected (22/37, 60%) compared with their counterparts undergoing uterine-sparing procedures (11/38, 29%; $P=.008$).

Patient-Reported Value of Text Messaging

The third and final question posed to the 62 survey respondents who received text messages during their postoperative survey was “How valuable did you find the text messages while recovering at home?” Of the patients receiving text messaging, 65% (40/62) reported value in the messaging (Figure 3).
**Discussion**

**Principal Findings**

Implementing a postoperative text messaging service is a feasible and potentially cost-effective way to deliver postoperative discharge instructions to patients undergoing laparoscopic benign gynecologic surgery. Text messaging was successfully received by 80% of intended intervention patients. This was accomplished at a cost of US $2.85 per patient. The study demonstrated patients receiving postoperative text messaging showed a trend toward an increase in satisfaction with their recovery and statistically significant increase in their sense of connection with their health care team. Largely, patients receiving text messaging found the texts valuable.

An unexpected finding was that patients undergoing uterine-sparing procedures had lower scores at baseline for satisfaction with recovery and connection with the health care team. These differences were eliminated with the text messaging intervention. A possible explanation may be the patients’ social support networks. Many women have undergone hysterectomies and are often willing to share advice on recovery, which might supplement the patient’s experience in recovery. Patients undergoing uterine-sparing procedures may not have as much access to this type of support. These findings may make the text message intervention even more valuable in a patient population undergoing uterine-sparing procedures.

Among patients selected to receive text messages, 20% did not receive the initial welcome text message, most likely due to use of a landline number instead of a mobile number and incorrect numbers in their charts. This error can be mitigated in a future study; however, this may limit eligible participants. The total cost for the text messages in this pilot study was US $178 to provide the full text message series to a total of 62 patients. It is difficult to quantify the financial benefit of increased patient satisfaction and connectedness. It is possible postoperative text messaging affords a significant return on investment through decreased complications requiring admission, decreased patient phone calls or messages, and decreased follow-up visits. Previous studies have demonstrated that the use of a mobile application for low-risk, postoperative, ambulatory patients led to deceased in-person follow-up visits [9]. The extension of this text messaging protocol to a larger number of patients may be cost-prohibitive if patients’ goodwill and a reduction in resource usage cannot be demonstrated. This study used a convenience sample of patients who received surgery over a 6-month period, so likely was not powered to detect statistically significant differences. A larger study that tracks complication rates and surveys office staff and providers is also necessary.

**Conclusions**

Postoperative text messages proved to be feasible in a population of patients undergoing benign gynecologic laparoscopic procedures. Text messages demonstrated a trend toward increased patient satisfaction with recovery and statistically significant increase in perception of connection with the health care team. The trend was more pronounced in patients undergoing uterine-sparing procedures. Minimal errors in messaging were identified. Given increasing emphasis on patient experience and the practice of cost-effective health care, further evaluation of a postoperative text messaging protocol that is adequately powered is warranted to determine patient and resource allocation benefit.

**Acknowledgments**

Funding was received through departmental research funds.
Conflicts of Interest
None declared.

References
Abstract

Many promising telemedicine innovations fail to be accepted and used over time, and there are longstanding questions about how to best evaluate telemedicine services and other health information technologies. In response to these challenges, there is a growing interest in how to take the sociotechnical complexity of health care into account during design, implementation, and evaluation. This paper discusses the methodological implications of this complexity and how the sociotechnical context holds the key to understanding the effects and outcomes of telemedicine. Examples from a work domain analysis of a surgical setting, where a telemedicine service for remote surgical consultation was to be introduced, are used to show how abstracted functional modeling can provide a structured and rigorous means to analyze and represent the implementation context in complex health care settings.

KEYWORDS
telemedicine; telemedicine evaluation; ERCP; work domain analysis; abstraction hierarchy; complexity; context; cognitive systems engineering

Introduction

Overview

Time has shown that it is difficult to scale up successful telemedicine innovations, and telemedicine has long been fraught with critical dilemmas regarding implementation, adoption, and evaluation [1-4]. In order to move forward from the repeated and seemingly paradoxical failures of telemedicine [2] and health information technologies, there have been calls for research and design methodologies that can address the many levels of complexity in health and care [5,6]. In this paper, we present reasoning for why it is important to apply a “complexity lens” to understand baseline conditions prior to technology implementation and evaluation in health care settings, and a rationale for why mapping the context provides important keys to understanding clinical outcomes and adoption when new health technology interventions are introduced. We describe how principles from complexity science can be applied in a structured and rigorous analysis of a telemedicine implementation context through work domain analysis [7-9]. Work domain analysis is a type of modeling specifically developed to design and analyze complex, adaptive sociotechnical systems. We include examples of how the method was used to analyze and represent many different sources of complexity that shape work in a surgical setting [10].

Context and Complexity in Health Technology Implementation and Evaluation

It is generally acknowledged that health technology implementation and outcomes are affected by contextual factors, and it is extremely difficult to scale up demonstration projects [5]. Despite this, few studies account for the preconditions for implementation in a way that adequately captures the inherent complexity of health care or in a fashion that can inform systems development or assessment. Technological interventions in
health care are generally complex, as are the health care settings in which they are used, and there is a demand for a methodological shift when studying the introduction of new technologies in this type of context [11-13].

We encountered a number of challenges when we attempted to provide a baseline description of the implementation context when a telemedicine system for remote surgical consultation was to be scaled up to multiple hospitals. If the service was adopted and used over time, it was expected to improve clinical outcomes and provide educational benefits. However, there were differences between the hospitals (eg, in work practices and resources), which potentially could interact with adoption and even lead to abandonment [5]. In addition to the inherent complexity of the clinical procedure and patients’ conditions, introducing new technology for surgical collaboration introduced new sources of complexity; for instance, the telemedicine system would bridge several technical systems and social and professional workgroups at different hospitals. These factors together would contribute to system-level outcomes over time (Figure 1).

Figure 1. System level outcomes. ERCP: endoscopic retrograde cholangio-pancreatography.

It was unclear to us how technical, social, and organizational factors set the implementation sites apart, how to conduct the analysis, and how to represent the findings in a way that could be useful for stakeholders.

A systems engineering method called work domain analysis (WDA) [7-9] appeared to be a suitable method for our purposes. WDA is intended to analyze complex work systems, and there are examples where it has been used in health care [14-19]. However, in our case, we wanted to capture and contrast domain-specific contextual factors on multiple work system levels at multiple sites, and this broad, explorative scope of analysis presented a particular modeling challenge. We eventually found a solution to this dilemma through iterative modeling, whereby we found a way to construct multiple, complementing models of the domain.

In the following sections, we provide a background for our choice of method.

**Evaluating Telemedicine**

There are many reasons for evaluating new technologies in health care and many ways to do this. While it might seem evident that new telemedicine systems should be evaluated for clinical, policy, and economic reasons, this is not always the case. In 2010, only 20% of the WHO (World Health Organization) members reported having published an evaluation or review on the use of telemedicine in 2006 [20]. A review of evaluations studying deployed hospital-to-hospital telemedicine services up to May 2016 only identified 164 papers [21].

Telemedicine evaluation raises many research questions [22], and different stakeholders have different expectations about what an evaluation should provide: evaluations can include organizational, technical, social, ethical, and legal, as well as transferability aspects [23]. Among peer-reviewed clinical evaluations of telemedicine interventions in hospital facilities, half were evaluated in terms of clinical outcomes and economic or satisfaction measures, while the other half were descriptive reports with ad hoc structure [21].

**Telemedicine as an Intervention**

In clinically-oriented research, telemedicine can be described as an intervention, which emphasizes its function in a clinical process [24]. Evaluating an intervention’s effectiveness and efficacy is central to health care quality since practice recommendations should be grounded in high-quality evidence, and policy and funding decisions should also have a sufficient basis [25]. Randomized controlled trials are a gold standard for ascertaining the efficacy of pharmaceuticals or clinical treatments with well-defined active components. This type of study is designed to show the size of a clinically meaningful benefit and the likelihood that this result is caused by the intervention (ie, that the intervention causes an effect X with size Y, with a confidence interval of Z). In addition, the results can show that an intervention is safe and effective [26].

Controlled studies are feasible for interventions with a limited number of readily defined components and a known mechanism. However, in the case of complex interventions [27] such as telemedicine [28], the “active components” can be difficult to define, and it might not be entirely clear what changes can be expected or how change will be achieved [29].

In response to the challenges of evaluating complex interventions, there is guidance recommending that process evaluations be conducted alongside trials of complex interventions to help provide the information required to
interpret the findings of a specific evaluation and to generalize findings to other contexts [29]. Three major themes in process evaluation are implementation, mechanisms, and context [27].

Implementation, Mechanisms, and Context

Implementation research has identified and compiled a multitude of contextual factors that can influence how new treatments and new ways of working are accepted and adopted [30,31,34]. The profusion of contextual factors that can affect an intervention reflects the reality of health care settings, where there is generally very much going on. Guidance for evaluating complex interventions also recommends that the intervention and its assumed causal mechanisms be clearly described, but this is difficult to achieve without a systematic description of the implementation context [32].

While context is considered to be a major concern for outcomes of complex interventions, the term remains “a slippery notion” [33], which is inconsistently defined and conceptualized [30,31,34]. In complex, adaptive work systems, which per definition are intractable, it is untenable to identify and measure all potential determinants [35]. The “variables paradigm” [36] provides output such as lists and categorizations of key variables and evaluation elements, yet is unlikely to provide means to identify which possible factors are most likely to influence outcomes in a particular case or how these factors interact once an intervention is introduced [33].

In-depth case research using sociological and organizational research methodologies contrasts to predominant “mechanistic” conceptions about implementation and component-oriented research [37] and can provide keys to the mechanisms of implementation and adoption through detailed experiential and contextual information [38,39]. The internal validity of such approaches is gained by the authenticity of observations and interpretations, which can cause weaknesses (eg, in quality case reports where a few members recount their experience of an intervention and its impact) [40]. In both case study research and quality reports, results are often presented in the form of detailed accounts, interspersed with quotes from participants [40]. This format can capture the uniqueness of a case but simultaneously raises questions about what has been accounted for, how lessons can be transferred to other interventions, or how the same intervention may have played out in another setting.

Common trial reporting formats and health technology assessments generally demand limited information about an intervention’s implementation or context [41]. Yet without this information, generalizing findings beyond a specific case becomes difficult. It also becomes challenging to determine whether changes detected during a study are due to the intervention or if its implementation or context is causing the effects [26].

These insights have generated calls to apply concepts from complexity theory during implementation and evaluation [6,13,29,42].

Complexity in Health Care

In health services research, “complexity science” has been used as an umbrella term, referring to the use of concepts about complex adaptive systems as a response to the increasing complexity and rapid rate of change and of health care [43,44]. “Complexity science” may invoke associations to definitions and methods used to address computational complexity and natural systems, and its use in health services research has invoked some criticism from proponents of “hard” approaches to complexity [45]. However, complexity science ideas have been, for example, used to inform theories and frameworks for evidence translation, implementation, and evaluation [30,39,46,47].

Thus far, health services research has mainly used complexity concepts to “sensitize” and support evaluation and also help identify issues that need to be managed during implementation, for example, in workshops to identify or solve specific problems, increase collaborative practices, or identify barriers to change [48]. However, complexity concepts have been inconsistently applied [48] or merely used as an abstract explanatory tool when they can be disciplined and refined to match specific research questions [49]. Moreover, the superficial use of complexity concepts runs the risk of fixating on easily identified components of a system, or effects from the context, without breaking adaptations apart or systematically taking interactions into account [50,51]. Another effect can be that different levels of work systems are separated and studied by different disciplines [52] or approached through different studies [47].

However, there is an option to apply systems engineering and research methodologies developed specifically for complex settings. Within the field of human factors or ergonomics, there is a long history of employing systems approaches in research and design of human-technology interactions [53], with concepts and definitions that explicitly address behavioral and organizational factors. The field provides theory and engineering methodologies employed in regulated, high-performance, and safety-critical domains such as aviation, the nuclear industry, and defense, but which are also well-suited for health care [54-56]. These types of work systems share characteristics; work is conducted by humans and technology, with operators balancing performance, quality, and safety with the demands set by uncertainty and rapid technological and organizational change.

Complex, Adaptive Sociotechnical Systems

Health care settings such as hospitals can be defined as complex, adaptive sociotechnical systems [57,58], where interactions among technical, human, and organizational elements generate complexity in many dimensions. This implies that efforts to induce system change through new technology can be expected to affect patterns of interaction within the system and between the system and its context [59,60]. These patterns of interactions make it challenging to scale up innovations from one context to another [61] and also make evaluation difficult, as it is hard to link technological change to specific outcomes [62].

One value of a sociotechnical systems approach lies in acknowledging the variable and irregular nature of health care
work, where staff continually monitor and adapt to changing circumstances to uphold safety and performance while helping patients and balancing organizational demands, social and professional values, and norms [63]. This intentional, adaptive behavior generates the system’s “self-organizing” capacity, making these work systems resilient [64] and providing keys to understanding and describing the system.

Principles from theory about complex adaptive systems emphasize the need to understand initial conditions as a baseline, and a “map” of the context and the rules governing behavior are important for understanding system behavior [6,65].

Cognitive Systems Engineering

Cognitive systems engineering (CSE) is a field of research and design in complex, adaptive sociotechnical systems [66-68]. CSE research focuses on how designed artifacts interact with their environment and with the humans using it. CSE practice includes eliciting and defining requirements and designing and evaluating human-technology work systems [69].

Cognitive work analysis (CWA) [8,70] is a formative CSE framework that provides functional analysis methods for design and evaluation. CWA is rooted in traditions and concepts from ecological psychology [71], distributed cognition [72], and expert decision-making in naturalistic settings [73], which emphasize how mutual interactions between the environment and agents shape behavior.

CWA consists of five phases, the first of which is called WDA [8,9]. WDAs are typically performed to provide a shared representation of complex work systems in the face of technological change ahead of system development or acquisition (eg, during the design requirements and specifications phases) [9]. WDA focuses on modeling the contextual factors which shape actors’ behaviors. The idea is to create a complete picture of the workers’ problem space by specifying what is to be achieved and the values, priorities, functions, and physical resources that affect this work.

This is done in abstracted, functional terms rather than through details of objects or tasks. Representing the sociotechnical context in this abstracted format can support the exploration of how the affordances of physical objects interact with functions towards system goals and how expanding the system with new components may impact the system as a whole [74].

Can Work Domain Analysis be Applied in Health Care Work Systems?

CWA originated in analyses of well-defined, tightly coupled causal systems (ie, engineered systems that are constrained by natural laws and technical factors) [7]. Some have claimed that WDA is not well-suited for health care [75], where system behavior is characterized by intentional constraints, such as actors’ goals, values, priorities, and shared rules of practice.

Health care systems are generally open systems, with many external interactions. This means that it is difficult to define system boundaries and distinguish discrete components and mechanisms that are “internal” to the system and how they interact with the “outer environment” [76]. As a consequence, boundaries between what is internal and what is external will be conceptual, an artifact, rather than ontological [8], and must be decided with careful consideration of the purpose of the analysis [9].

However, WDA has been used to model numerous intentional, open systems(eg, in naval command and control, ambulance dispatch, and health care) [77-79]. Jenkins et al suggest that if suitably adapted, WDA models can be utilized to predict and evaluate system-level outcomes when new technologies are introduced in sociotechnical systems [62].

Understanding the Implementation Context for Teleguidance in ERCP

The telemicine service we studied was developed to enable real-time, professional-to-professional video collaboration during endoscopic retrograde cholangio-pancreatography (ERCP), a technically advanced endoscopic procedure for biliary and pancreatic disease. The telemicine service, which came to be called teleguidance, had demonstrated clinical and economic benefits in a feasibility study [80]. Health-economic modeling also showed the potential for positive clinical and economic outcomes [81]. This provided a rationale for scaling up the practice and an interest to generate additional evidence for the new way of working.

Quantitative clinical data was to be collected to investigate the clinical effectiveness of teleguidance. However, the service also had to be used over time for any desired quality improvement outcomes to come into effect. So there was also interest to conduct a qualitative inquiry to understand whether conditions at the various sites might influence any clinical results or affect how teleguidance would be adopted and assimilated into everyday practice (Multimedia Appendix 1).

The complexity of the highly specialized surgical procedure and the hospital settings made it neither feasible nor theoretically justified to choose an ad hoc number of components (eg, from a determinant framework) [31], and also expect to achieve an adequate understanding of how the implementation context could influence the use and outcomes of the new technology [11]. Without a method adapted for the complexity of the work systems involved, our attempt at understanding and describing the implementation context might also fail to account for interactions between different parts of the system, or from the context, in a systematic way.

We observed WDA was seen as a candidate method as it is developed to accommodate sociotechnical complexity and would provide a structured and accountable analysis. Representing constraints that shape system behavior in an abstracted manner would be useful for comparing different sites and could also support the prediction of change and unintended consequences, which is a central aspect of complexity-informed evaluation [6].

Methods

Data Collection and Analysis

We used an ethnographical approach with extensive fieldwork and interviews to collect data and generate a deep understanding of the context in a working system [82]. This included three
iterations of data collection using a sequence of techniques, moving from a general “rough” level of description and understanding to a finer grain.

Observations and interviews were conducted at the central and remote sites. A total of 20 semistructured interviews with 10 ERCP specialists, 5 ERCP assistants, 3 technical staff, and 2 administrative staff from 5 hospitals were conducted. During the data collection phase, a service blueprint [83] was designed, which served as an intermediary, shared representation [84] (Multimedia Appendix 2). Interviews were transcribed, and thematic coding [85] was conducted, using predetermined categories and prompts (Multimedia Appendix 3) pertaining to the abstraction levels described by Naikar [9] from the WDA framework.

Data collection and analysis are described in more detail in an adjacent paper, “Modeling implementation context in telemedicine” [10].

Defining System Boundaries

We wanted to conduct a broad and deep investigation of the ERCP work context at the participating sites. Our broad interpretation of the implementation context was based on a view that implementation and use of teleguidance will be shaped by physical and organizational constraints, which are situated within the functional goals and values of the work domain [86].

When analyzing a system, it is generally deemed necessary to define system boundaries in order to distinguish components and mechanisms that are “internal” to the system, how they interact with the “outer environment,” and their functional relationships [78]. However, work system complexity became apparent early on during observations and interviews, and it became clear that it would be difficult to set clearly defined system boundaries.

During procedures, the ERCP team works within a specific physical space to perform a specific type of procedure during a limited time frame. However, before and during each ERCP procedure, there were continual trade-offs between clinical work and organizational demands (eg, demands for resource efficiency), which also sometimes conflict with clinical priorities. It also became clear that administrative and clinical roles and tasks were highly interwoven in ways that were not necessarily reflected by formal roles or organizational boundaries. Similarly, development work, such as research and training, was continually ongoing, and these aspects of work were shaped by other sets of constraints (eg, funding and practice standards, which were controlled by sources other than the hospital administration).

In addition, the clinical work system at the university hospital underwent substantial reorganization during the study. Similar but less extensive reorganizations were taking place at several of the smaller hospitals.

Hollnagel reframes the question of system boundaries and context by speaking in terms of foreground and background functions, rather than strict system boundaries [87]; as such, functions and constraints which are central during ERCP procedures could be included in the WDA model without explicit reference to whether they lay within an arbitrary system definition or not.

Creating the Models

The abstraction hierarchy was constructed using Naikar [9] as a main resource and iteratively modeling our findings and revisiting the purpose of the analysis, with feedback from clinical practitioners and project managers.

The first iteration of the abstraction hierarchy centered on structuring the large amounts of data collected. Findings from interviews, observations, and documents were entered into a large general-purpose spreadsheet. The findings were categorized into abstraction levels and linked to cells in the matrix.

Multimedia Appendix 4 shows the sequence in which we populated different levels of the abstraction hierarchy matrix and provides examples of the questions used to guide the work.

The second iteration of the abstraction hierarchy mainly focused on testing different ways of ranging and decomposing the clinical work system to make the analysis tractable. The interactions between clinical work and organizational demands made it challenging to define a part-whole systems decomposition of the hospitals and their subsystems according to organizational boundaries. We decided to represent “hospital” as the overarching system, with a partial abstraction hierarchy, and “ERCP work” as a conceptual subsystem, with a detailed abstraction hierarchy. The ERCP work system is loosely bounded [78], meaning ERCP practitioners have control over some resources but not of the whole system. Capturing this aspect was a conundrum until we decided to create multiple models of the domain. Naikar [88] proposes representing certain domains as having distinct facets. This is considered a way to handle the wide range of constraints necessary for this type of intentional system.

Therefore, a decision was made to construct multiple models of the domain, one representing the “primary” clinical work and the other representing the “secondary” functions that provide the infrastructure and resources for the clinical work, such as administration and management and training and research. The three facets represent aspects of the same clinical work system, yet each facet is seen as separate through the nature of tasks and aspects, such as organizational departments, competencies, and roles. Individual stakeholders can be involved with more than one of the three facets, as is the case with senior doctors and nurses who have managerial roles in addition to their clinical functions.

We considered development work (research, education, and training) to be distinct from other secondary functions, and we finally represented the domain as three facets: treatment, development, and administration.

The third iteration was an exercise in improving the internal structure of the means-ends relationships within the conceptual framework of the functional facets. The constraints were decomposed in detail on certain levels of abstraction but are aggregated in the presented model for increased legibility.
Using the Models to Proactively Identify Implementation Issues and Possible Outcomes

By tracing the many-to-many means-ends relationships among constraints through “how-what-why” reasoning [74], the abstraction hierarchies served as a simple artifact to investigate possible scenarios when teleguidance is used (Multimedia Appendix 5).

Results

Three functional facets of the domain were modeled (Figure 2) and are defined as follows:

- Clinical: The treatment of patients through ERCP (Multimedia Appendix 6).
- Development: Functions such as developing clinical methods and technology; research (eg, developing clinical methods and tracking outcomes), collaboration with suppliers, and arranging and providing supervision and training opportunities (Multimedia Appendix 7).
- Administration: Support functions for the clinical work, such as managing finances and staff and facilities, including IT (information technology) and medical technology (Multimedia Appendix 8).

The models helped us structure findings from the interviews, and we could link the physical teleguidance equipment to reverberations in logistic processes, such as patient and staff scheduling and preparations for ERCP procedures (eg, set-up, preparation of supplies, and team composition; Figure 3).

The abstraction hierarchies enabled us to explore possible scenarios during and after implementation. For example, we could identify issues that might be of importance during implementation and be weighed in during evaluation, such as if teleguidance affects the duration of procedures or the time to prepare for procedures.

A detailed account of specific findings is provided in “Modeling implementation context in telemedicine” [10].
Discussion

Overview

Despite general agreement that context is important and that “complexity science” can be of value, it appears that in practice, it is difficult to describe and analyze contextual factors and, at the same time, accommodate complexity during evaluation. Furthermore, context is often given a minor role in studies of health technology implementation or health technology assessment [30]. We decided to attempt a broad analysis of the implementation context for a telemedicine service by conducting a WDA. However, it was initially unclear if the scope and open, intentional nature of the work systems would be a problem.

Principal Findings

In order to represent the entire problem space that clinical practitioners face during ERCP, we discovered that it was relevant to include aspects of nonclinical work to the extent that they affected clinical procedures. To handle the width and depth of this scope, we conceptualized and modeled three functional “facets” of the domain which shape ERCP work (Figure 4).

Figure 4. Three functional facets of the domain which shape ERCP work. ERCP: endoscopic retrograde cholangio-pancreatograhy.

Each facet represents a set of constraints that shape ERCP team members’ work before, during, and after each ERCP procedure. With Hollnagel’s terminology [87], the functions and constraints within the administrative and development facets can loosely be considered as background to the functions and constraints in the clinical facet.

The models helped us explore the dynamics of the work systems and project possible interactions during the use of the telemedicine service. An example of findings was that despite shared clinical goals across the collaborating sites, relevant aspects of the administrative and development facets need similar coordination across hospitals. While this may seem obvious, these factors are beyond the control of the clinical staff and may interfere with teleguidance use over time. The WDA helped us identify and represent these types of issues in detail [10].

Defining System Boundaries

Initially, we set wide system boundaries to include many explanatory variables. During the data collection phase, it
became increasingly difficult to establish the boundaries for the work system and decide what should be included in the WDA. A narrow definition of the unit of analysis, for example, ranging the system according to what goes on within the physical space of the operation theater during ERCP, would have given a precise ontological boundary. However, this could have excluded organizational factors that have a bearing on ERCP performance and would therefore forfeit the purpose of the analysis, which was to map constraints that could come to interact with the implementation of teleguidance.

**Practical Aspects of Conducting a WDA**

The design, implementation, and assessment of the telemicine service we evaluated was a transdisciplinary effort involving clinical practitioners and researchers, alongside human-computer interaction and project management experts. Consequently, there were different expectations about what an evaluation should provide, and the practical application of WDA requires an understanding of systems-theoretical concepts. While Rasmussen [7,70] and Vicente [8] provide comprehensive but somewhat conflicting descriptions of how to conduct WDA, Naikar provides a systematic methodology [9]. However, it was challenging at first to construct an abstraction hierarchy due to the width of the analysis. We also had difficulties establishing a hierarchical decomposition of the work systems, as work in practice did not necessarily follow organizational boundaries.

Technical and physical factors are generally more straightforward to distinguish than properties emerging from human intentions. The qualitative methods we used for our deepening investigation required relationships between researchers and domain practitioners, which may be a hurdle due to interprofessional dynamics and hierarchies in health care and time constraints [89].

However, we conclude that the methodology is resource-efficient, especially if the analysis can be reused across multiple problems [90]. The structured, abstracted format is very compact and relatively easy for stakeholders from different disciplines, such as clinical staff and project management, to understand. WDA may be more useful in systems development and evaluations than the narrative accounts common in many qualitative case studies and thereby also be an effective artifact for supporting the interdisciplinary collaboration required for successful human-systems integration [84].

**Conclusion**

WDA is a systems engineering method that allowed us to create representations that served as objective models of the implementation context by focusing on functions and constraints shaping work-system behavior. Creating models helped us avoid the notion that context is a fixed entity or can be described by compiling variables or events. The three sets of constraints or facets, which were present in each hospital, represent constraints that shape everyday ERCP work and that can shape the use of teleguidance.

Using abstracted functional modeling guided by theory strengthens the transferability of findings, and the facets can be expected to reflect fields of interest and functions that can affect other telemedicine interventions in similar hospital settings. The structure of the method also supported an iterative “discovery and modeling” approach [91], which was necessary as our understanding of the work systems developed.

We conclude that WDA is an effective method for modeling the implementation context, and that this type of modeling is a practical approach to applying “complexity science” principles, and that it is a way to provide structured analysis without reducing complexity or detailed qualitative accounts common in sociological and organizational research methodologies. The models account for technological, social, and organizational factors and their dynamic interactions, which provide useful information both for policymakers and scientists. The method can be useful for supporting detailed analysis and planning prior to implementation and evaluation of telemedicine, which is currently rare [92].

**Future Work**

The three functional facets of the domain (clinical, development, and administration) represent sets of generic constraints that we believe are likely to be present in other hospital environments and likely to affect other technology implementation projects. These can serve as “dimensions” along which to model and analyze similar clinical work systems.

**Conflicts of Interest**

None declared

---

Multimedia Appendix 1

Aims of the clinical study that we were to complement with understanding of the implementation context.

[PNG File, 100 KB - periop_v4i2e26580_app1.png]

Multimedia Appendix 2

Service blueprint.

[PNG File, 319 KB - periop_v4i2e26580_app2.png]

Multimedia Appendix 3

Coding categories and criteria for the qualitative data analysis.

[PNG File, 28 KB - periop_v4i2e26580_app3.png]
Multimedia Appendix 4
The sequence in which we populated different levels of the abstraction hierarchy matrix.

[PNG File, 71 KB - periop_v4i2e26580_app4.png]

Multimedia Appendix 5
The abstraction hierarchies served as a simple artifact to investigate possible scenarios when teleguidance is used.

[PNG File, 69 KB - periop_v4i2e26580_app5.png]

Multimedia Appendix 6
The treatment facet.

[PNG File, 118 KB - periop_v4i2e26580_app6.png]

Multimedia Appendix 7
The development facet.

[PNG File, 146 KB - periop_v4i2e26580_app7.png]

Multimedia Appendix 8
The administrative facet.

[PNG File, 100 KB - periop_v4i2e26580_app8.png]

References


Abbreviations

- **CSE**: cognitive systems engineering
- **CWA**: cognitive work analysis
- **ERCP**: endoscopic retrograde cholangio-pancreatografy
- **WDA**: work domain analysis

©Hedvig Aminoff, Sebastiaan Meijer. Originally published in JMIR Perioperative Medicine (http://periop.jmir.org), 16.09.2021. 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.
Sustaining a Multidisciplinary, Single-Institution, Postoperative Mobilization Clinical Practice Improvement Program Following Hepatopancreatobiliary Surgery During the COVID-19 Pandemic: Prospective Cohort Study

Kai Siang Chan¹, MBBS; Bei Wang¹, MHA; Yen Pin Tan¹, MBBS, FRCS; Jaclyn Jie Ling Chow¹, MSc; Ee Ling Ong², BEng (Hons); Sameer P Junnarkar¹, FRCS, MD; Jee Keem Low¹, MBChB, MD, FRCS; Cheong Wei Terence Huey¹, MBBS, MRCS, MMed, FRCS; Vishal G Shelat¹, MBBS, MS, DNB, FRCS, FAMS, MCI

¹Department of General Surgery, Tan Tock Seng Hospital, Singapore, Singapore
²Office of Clinical Governance, Tan Tock Seng Hospital, Singapore, Singapore

Corresponding Author:
Vishal G Shelat, MBBS, MS, DNB, FRCS, FAMS, MCI
Department of General Surgery
Tan Tock Seng Hospital
11 Jalan Tan Tock Seng
Singapore, 308433
Singapore
Phone: 65 63577807
Email: vishal_g_shelat@ttsh.com.sg

Abstract

Background: The Enhanced Recovery After Surgery (ERAS) protocol has been recently extended to hepatopancreatobiliary (HPB) surgery, with excellent outcomes reported. Early mobilization is an essential facet of the ERAS protocol, but compliance has been reported to be poor. We recently reported our success in a 6-month clinical practice improvement program (CPIP) for early postoperative mobilization. During the COVID-19 pandemic, we experienced reduced staffing and resource availability, which can make CPIP sustainability difficult.

Objective: We report outcomes at 1 year following the implementation of our CPIP to improve postoperative mobilization in patients undergoing major HPB surgery during the COVID-19 pandemic.

Methods: We divided our study into 4 phases—phase 1: before CPIP implementation (January to April 2019); phase 2: CPIP implementation (May to September 2019); phase 3: post–CPIP implementation but prior to the COVID-19 pandemic (October 2019 to March 2020); and phase 4: post–CPIP implementation and during the pandemic (April 2020 to September 2020). Major HPB surgery was defined as any surgery on the liver, pancreas, and biliary system with a duration of >2 hours and with an anticipated blood loss of ≥500 ml. Study variables included length of hospital stay, distance ambulated on postoperative day (POD) 2, morbidity, balance measures (incidence of fall and accidental dislodgement of drains), and reasons for failure to achieve targets. Successful mobilization was defined as the ability to sit out of bed for >6 hours on POD 1 and ambulate ≥30 m on POD 2. The target mobilization rate was ≥75%.

Results: A total of 114 patients underwent major HPB surgery from phases 2 to 4 of our study, with 33 (29.0%), 45 (39.5%), and 36 (31.6%) patients in phases 2, 3, and 4, respectively. No baseline patient demographic data were collected for phase 1 (pre–CPIP implementation). The majority of the patients were male (n=79, 69.3%) and underwent hepatic surgery (n=92, 80.7%). A total of 76 (66.7%) patients underwent ON-Q PainBuster insertion intraoperatively. The median mobilization rate was 22% for phase 1, 78% for phases 2 and 3 combined, and 79% for phase 4. The mean pain score was 2.7 (SD 1.0) on POD 1 and 1.8 (SD 1.5) on POD 2. The median length of hospitalization was 6 days (IQR 5-11.8). There were no falls or accidental dislodgement of drains. Six patients (5.3%) had pneumonia, and 21 (18.4%) patients failed to ambulate ≥30 m on POD 2 from phases 2 to 4. The most common reason for failure to achieve the ambulation target was pain (6/21, 28.6%) and lethargy or giddiness (5/21, 23.8%).
Conclusions: This follow-up study demonstrates the sustainability of our CPIP in improving early postoperative mobilization rates following major HPB surgery 1 year after implementation, even during the COVID-19 pandemic. Further large-scale, multi-institutional prospective studies should be conducted to assess compliance and determine its sustainability.


KEYWORDS
enhanced recovery after surgery; early mobilization; liver resection; pancreatoduodenectomy

Introduction

Enhanced Recovery After Surgery (ERAS) is a multimodal, multidisciplinary perioperative approach to improve surgical outcomes [1]. The implementation of ERAS has improved perioperative outcomes in patients undergoing elective major hepatopancreatobiliary (HPB) surgery [2]. Postoperative early mobilization is an integral component of the ERAS protocols as it reduces pleuropulmonary complications and deep vein thrombosis [3]. Early postoperative mobilization also reduces postoperative ileus and length of hospital stay [4,5]. However, there are no standardized criteria to define mobilization, and compliance remains poor. Vague terminologies, including sitting out of bed, standing at the bedside, walking duration, and walking distances, are used to define mobilization. Recently, Grass et al [6] performed a retrospective study involving 1170 patients who had colorectal surgery in Switzerland to assess early postoperative mobilization (defined as sitting out of bed ≥6 hours on postoperative day [POD] 1). They showed that 58% of patients were noncompliant, with resulting increased postoperative morbidity (overall complications 55% vs 29%, P<.001) and length of stay (mean 12, SD 14 days vs mean 6, SD 7 days; P<.001) compared to the early mobilization group [6].

A systematic review by Coolsen et al [2] in 2013 described poor compliance (mobilization rate 20%-28%) to early postoperative mobilization on POD 1 following liver surgery [7,8]. Similarly, our institution showed a poor postoperative mobilization rate of 22% in patients undergoing elective major HPB surgery, with improvement to >75% following the implementation of a multidisciplinary surgeon-led clinical practice improvement project (CPIP) [9]. The quality improvement process does not end with the implementation of a solution. Specific steps must be taken, and mechanisms established to hold the gains, for breakthroughs in results come from sustaining changes. The Royal College of Physicians of London, United Kingdom, has incorporated sustainability within the Institute of Medicine’s six quality domains [10]. The median follow-up time for a health care CPIP is reported to be less than 1 year [11]. Only a sustained initiative can be spread for adoption by others at multiple locations so that communities can reap gains.

Ensuring sustainability is difficult due to the COVID-19 pandemic. The COVID-19 pandemic has had a profound impact on the community, health care workers, and health care systems, with more than 3.1 million deaths as of May 2021 [12]. In light of this pandemic, our institution reallocated resources preferentially for COVID-19–related care to cope with clinical demands. Our HPB unit began triaging and scaling down elective surgery to facilitate staff redeployment and reduce patient exposure to the novel coronavirus. Oncology-related services, however, were minimally disrupted given the time-sensitive nature of these diseases and the need for prompt management [13]. Saab et al [14] surveyed 82 centers in 28 countries and described reduced pain management and supportive care services by 26% and limitations in social services support by 74%. To add on, mobilization mandates staff to be near patients, which violates safe distancing measures. A clinical practice guideline by Thomas et al [15] in 2020 for physiotherapy management during the COVID-19 pandemic recommended screening referrals for mobilization and exercise to minimize staff in contact and high-filtration masks during physiotherapy sessions. Locally, personal protective equipment was mandatory for physiotherapists, and ambulation was limited to the patients’ ward cubicile to minimize external contact. There are also increased stressors associated with fear and anxiety of becoming infected [16]. Hence, this study aimed to assess the sustainability of our multidisciplinary single-institution CPIP at 1-year postimplementation to improve the postoperative mobilization rate of patients undergoing elective major HPB surgery during the COVID-19 pandemic.

Methods

Overview

Our institution is a university-affiliated tertiary hospital with 1700 inpatient beds. ERAS started in March 2016 in the colorectal surgery division. In line with the concept of ERAS, the HPB unit began inpatient prehabilitation for patients undergoing elective liver surgery in 2016, resulting in a reduction in overall morbidity and improved social well-being [17]. The entire ERAS protocol subsequently expanded to the HPB unit for patients undergoing elective major HPB surgery in 2018. The HPB surgery dashboard for 2018 following the implementation of ERAS revealed a low observed/expected ratio for compliance, with a postoperative mobilization rate of 22%. Hence, relevant stakeholders agreed to implement a CPIP to improve postoperative mobilization, which began in May 2019 [9].

Study Protocol

The specific details of our CPIP were described in 2020 by Tang et al [9]. Successful mobilization was defined as sitting out of bed for ≥6 hours on POD 1 and ambulation of ≥30 m on POD 2, with a target mobilization rate of ≥75%. Preoperatively, case managers counsel patients and caregivers on postoperative goals and emphasize the benefits of early mobilization. Postoperatively, the surgical teams emphasize the benefits of
mobilization during POD 1 evening rounds. The plan-do-study-act (PDSA) cycles were utilized to identify critical barriers to early mobilization, and changes were implemented to identify outcomes. Major HPB surgery was defined as surgery involving the HPB system and lasting more than 2 hours. Patients were excluded from the study if they had undergone cholecystectomy, common bile duct exploration, laparotomy for general surgical conditions, or major HPB surgery with intraoperative blood loss of ≥2 L or a surgery duration of >9 hours.

We summarized the entire mobilization improvement process into 4 phases:

- Phase 1 (January to April 2019): prior to CPIP implementation;
- Phase 2 (May to September 2019): CPIP implementation, where there is direct oversight to improve postoperative mobilization using the PDSA cycles;
- Phase 3 (October 2019 to March 2020): post-CPIP, before the COVID-19 pandemic, where there was indirect oversight of postoperative mobilization. This was also required routinely as part of our institution’s protocol following CPIP implementation;
- Phase 4 (April 2020 to September 2020): post-CPIP, during the COVID-19 pandemic, where there was no oversight on interventions to improve postoperative mobilization.

Figure 1 is a schematic representation of the 4 phases of the mobilization improvement process. As the purpose of this study is to assess the sustainability following our CPIP during the COVID-19 pandemic, we will be primarily describing phase 4 of our study.

*Sustainability Run Chart: Mobilization Rates on POD 2 (Period: January 2019 to September 2020)*

<table>
<thead>
<tr>
<th>Phase 1 (Pre-CPIP)</th>
<th>Phase 2 (During CPIP)</th>
<th>Phase 3 (Before COVID-19)</th>
<th>Phase 4 (During COVID-19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Mobilization Rate</td>
<td>22%</td>
<td>78%</td>
<td>79%</td>
</tr>
<tr>
<td>Target = 75%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Impact of COVID-19 Locally

The first case of COVID-19 in Singapore was detected in January 2020, and the national “circuit breaker measures” were announced on April 7, 2020 [18]. This gave us the chance to stratify the analysis into phase 3 (before COVID-19, from October 2019 to March 2020) and phase 4 (during COVID-19, from April 2020 to September 2020). During phase 4, there...
were three modifications to postoperative mobilization due to COVID-19: (1) physiotherapists had to wear full personal protective equipment during physiotherapy sessions, (2) mobilization was limited to within the patients’ cubicles to limit the risk of infection, and (3) segregation of physiotherapy teams to reduce cross-contact among the health care personnel.

Perioperative Management and Prehabilitation Program

In 2016, the HPB unit introduced an inpatient prehabilitation program, 2 to 4 weeks in duration [17]. The program involves a multidisciplinary team comprising physiotherapists, dieticians for nutrition optimization, and case managers for patient education on the surgery and postoperative expectations. All elective major HPB surgical patients were considered for the program unless excluded due to logistic reasons or resource constraints (eg, the surgery date was too close or there was a lack of program slots). Pain management is an integral part of the prehabilitation program and an essential component relevant to mobilization. We adopted a multimodal approach to manage postoperative pain; the ON-Q PainBuster (B Braun Melsungen AG), an elastomeric pump device providing a continuous infusion of 400 ml of ropivacaine, was intraoperatively inserted at the discretion of the surgeon in the preperitoneal space. The ON-Q PainBuster was not routinely inserted for laparoscopic procedures. The majority of patients were also on patient-controlled analgesia (PCA) (fentanyl) and paracetamol postoperatively. Epidural analgesia is infrequently used at our institution in view of the following reasons: (1) placement and removal of epidural analgesia is more technically challenging and slower compared to the ON-Q PainBuster; (2) recommendations from the American Society of Regional Anesthesia that an international normalized ratio (INR) of 1.4 is the upper limit for safe removal of an epidural catheter (in our group of patients undergoing major hepatic resections, an abnormal INR is to be expected and would require fresh frozen plasma coverage [19]); and (3) local protocol mandating the need for monitoring in high-dependency units while on epidural analgesia (this precludes patients who are clinically improving and stable after transfer to the general ward and occupy limited high-dependency beds for patients who may require high-dependency monitoring).

Data Collection and Study Variables

Data were extracted for all patients included in this study from phases 2 to 4 (May 2019 to September 2020) from a prospectively maintained standing database for patients undergoing HPB surgery approved by the local institutional review board. Data were not extracted for patients for phase 1 (before CPIP implementation). Study variables included insertion of the ON-Q PainBuster intraoperatively, pain score, length of hospital stay, distance ambulated on POD 2, morbidity, balance measures, and reasons for failure to achieve targets. Length of hospitalization stay was defined as the duration of hospital stay calculated from admission to the point of discharge. Successful mobilization was defined as sitting out of bed for >6 hours on POD 1 and ambulation of ≥30 m on POD 2 [9]. Morbidity was defined as the incidence of pneumonia and deep vein thrombosis. Balance measures, defined as potential complications secondary to mobilization—incidence of falls and accidental drain dislodgement—were evaluated.

Statistical Analysis

All data were tabulated into an Excel sheet (Microsoft Corp) and transposed into SPSS, version 25.0 (IBM Corp), for statistical analysis. Categorical values were described as percentages and analyzed by the chi-square test and Fisher exact test for variables with expected cell count <5. Continuous variables were described as median (IQR) or mean (SD) and were analyzed by the Kruskal-Wallis test or analysis of variance (ANOVA), respectively. Statistical significance was defined as a P value of <.05.

Results

Baseline Demographics and Clinical Profile

A total of 114 patients underwent major HPB surgery from phases 2 to 4 of our study, with 33 (29.0%), 45 (39.5%), and 36 (31.6%) patients in phases 2, 3, and 4, respectively. Baseline patient demographic data were not collected for phase 1 (pre-CPIP implementation); information on the median mobilization rate of 22% during phase 1 was obtained from the HPB surgery dashboard in 2018 following the implementation of the ERAS protocol in our HPB unit. The majority of the patients were male (n=79, 69.3%) and underwent hepatic surgery (n=92, 80.7%). There were 76 (66.7%) patients who had an ON-Q PainBuster insertion intraoperatively. Table 1 summarizes the demographics of the study population from phases 2 to 4.
### Table 1. Patient demographics and clinical profile of the study population.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall cohort&lt;sup&gt;a&lt;/sup&gt; (n=114)</th>
<th>Phase 2&lt;sup&gt;b&lt;/sup&gt; (n=33)</th>
<th>Phase 3&lt;sup&gt;c&lt;/sup&gt; (n=45)</th>
<th>Phase 4&lt;sup&gt;d&lt;/sup&gt; (n=36)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), median (IQR)</td>
<td>66.5 (60.8-71.3)</td>
<td>67 (56-71)</td>
<td>66 (61-71)</td>
<td>67 (61.3-72)</td>
<td>.74</td>
</tr>
<tr>
<td>Gender (male), n (%)</td>
<td>79 (69.3)</td>
<td>20 (60.6)</td>
<td>31 (68.9)</td>
<td>28 (77.8)</td>
<td>.30</td>
</tr>
<tr>
<td>ASA score&lt;sup&gt;e&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.19</td>
</tr>
<tr>
<td>I</td>
<td>1 (0.9)</td>
<td>1 (3.0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>38 (33.3)</td>
<td>11 (33.3)</td>
<td>19 (42.2)</td>
<td>8 (22.2)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>75 (65.8)</td>
<td>21 (63.6)</td>
<td>26 (57.8)</td>
<td>28 (77.8)</td>
<td></td>
</tr>
<tr>
<td>Prehabilitation, n (%)</td>
<td>63 (55.3)</td>
<td>21 (63.6)</td>
<td>19 (42.2)</td>
<td>23 (63.9)</td>
<td>.08</td>
</tr>
<tr>
<td>Surgical approach, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.01</td>
</tr>
<tr>
<td>Laparoscopic</td>
<td>40 (35.1)</td>
<td>13 (39.4)</td>
<td>12 (26.7)</td>
<td>15 (41.7)</td>
<td></td>
</tr>
<tr>
<td>Laparoscopic converted open</td>
<td>5 (4.4)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>5 (13.9)</td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>69 (60.5)</td>
<td>20 (60.6)</td>
<td>33 (73.3)</td>
<td>16 (44.4)</td>
<td></td>
</tr>
<tr>
<td>Type of surgery, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.42</td>
</tr>
<tr>
<td>Hepatic</td>
<td>92 (80.7)</td>
<td>25 (75.8)</td>
<td>39 (86.7)</td>
<td>28 (77.8)</td>
<td></td>
</tr>
<tr>
<td>Pancreatic</td>
<td>22 (19.3)</td>
<td>9 (24.2)</td>
<td>6 (13.3)</td>
<td>8 (22.2)</td>
<td></td>
</tr>
<tr>
<td>Placement of ON-Q PainBuster (yes), n (%)</td>
<td>76 (66.7)</td>
<td>20 (60.6)</td>
<td>33 (73.3)</td>
<td>23 (63.9)</td>
<td>.46</td>
</tr>
<tr>
<td>Abdominal drains, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.49</td>
</tr>
<tr>
<td>0</td>
<td>31 (27.2)</td>
<td>12 (36.4)</td>
<td>13 (28.9)</td>
<td>6 (16.7)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>69 (60.5)</td>
<td>17 (51.5)</td>
<td>27 (60.0)</td>
<td>25 (69.4)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13 (11.4)</td>
<td>4 (12.1)</td>
<td>4 (8.9)</td>
<td>5 (13.9)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1 (0.9)</td>
<td>0 (0)</td>
<td>1 (2.2)</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Overall cohort refers to the study population from phases 2 to 4. Study demographics are not shown for phase 1 patients.  
<sup>b</sup>Phase 2 refers to the period during CPIP implementation (May 2019 to September 2019).  
<sup>c</sup>Phase 3 refers to the sustainability phase post-CPIP but before the COVID-19 pandemic (October 2019 to March 2020).  
<sup>d</sup>Phase 4 refers to the sustainability phase post-CPIP during the COVID-19 pandemic (April 2020 to September 2020).  
<sup>e</sup>ASA: American Society of Anesthesiologists.

### Postoperative Outcomes

Table 2 summarizes the outcomes of the study population in each phase of the study.

The median mobilization rate was 22% for phase 1, 78% for phases 2 and 3 combined, and 79% for phase 4 (Figure 1). We combined the median mobilization rate for phases 2 and 3 as the mobilization rate at the start of phase 2 will be lower since it takes time for the effects of the CPIP to be seen; in April 2019 (prior to the start of phase 2), mobilization was 2 out of 10 (20%). The mean pain score was 2.7 (SD 1.0) on POD 1 and 1.8 (SD 1.5) on POD 2. A pairwise comparison of pain score on POD 2 showed a significant difference in pain score between phases 2 and 4 (phase 2: pain score 2.3, SD 1.8 vs phase 4: pain score 1.3, SD 1.3; P=.01). The median length of hospital stay was 6 days (IQR 5-11.8). There were no falls or accidental dislodgement of drains. A total of 6 patients (5.3%) had pneumonia.
Table 2. Study population outcomes.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall cohort(^a) (n=114)</th>
<th>Phase 2(^b) (n=33)</th>
<th>Phase 3(^c) (n=45)</th>
<th>Phase 4(^d) (n=36)</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating time (min), median (IQR)</td>
<td>338 (240-489)</td>
<td>248 (178-379)</td>
<td>285 (228-353)</td>
<td>350 (259-456)</td>
<td>.01</td>
</tr>
<tr>
<td>Blood transfusion (yes), n (%)</td>
<td>20 (17.5)</td>
<td>5 (15.2)</td>
<td>6 (13.3)</td>
<td>9 (25.0)</td>
<td>.36</td>
</tr>
<tr>
<td>Length of hospital stay (days), median (IQR)</td>
<td>6 (5-11.8)</td>
<td>6 (4-8)</td>
<td>6 (5-8)</td>
<td>6.5 (4-17)</td>
<td>.79</td>
</tr>
<tr>
<td>Pain score, mean (SD)</td>
<td>2.7 (1.0)</td>
<td>2.7 (0.7)</td>
<td>2.7 (1.0)</td>
<td>2.7 (1.5)</td>
<td>.98</td>
</tr>
<tr>
<td>Ambulated (\geq 30) m on POD 2 (yes), n (%)</td>
<td>93 (81.6)</td>
<td>24 (72.7)</td>
<td>40 (88.9)</td>
<td>29 (80.6)</td>
<td>.19</td>
</tr>
<tr>
<td>Actual distance walked (m), median (IQR)</td>
<td>50 (30-100)</td>
<td>40 (21-100)</td>
<td>70 (45-100)</td>
<td>50 (30-100)</td>
<td>.16</td>
</tr>
<tr>
<td>Reasons for failing to achieve target ambulation (n=21), n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>6 (28.6)</td>
<td>2 (22.2)</td>
<td>2 (40.0)</td>
<td>2 (28.6)</td>
<td>.51</td>
</tr>
<tr>
<td>Lethargy/giddiness</td>
<td>5 (23.8)</td>
<td>3 (33.3)</td>
<td>1 (20.0)</td>
<td>1 (14.3)</td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td>2 (9.5)</td>
<td>1 (11.1)</td>
<td>1 (20.0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Hypotension/tachycardia</td>
<td>2 (9.5)</td>
<td>1 (11.1)</td>
<td>0 (0)</td>
<td>1 (14.3)</td>
<td></td>
</tr>
<tr>
<td>Medical instructions (postchest tube removal)</td>
<td>2 (9.5)</td>
<td>2 (22.2)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Local protocols (ongoing blood transfusion)</td>
<td>2 (9.5)</td>
<td>0 (0)</td>
<td>1 (20.0)</td>
<td>1 (14.3)</td>
<td></td>
</tr>
<tr>
<td>Admitted to ICU(^e)</td>
<td>2 (9.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (28.6)</td>
<td></td>
</tr>
<tr>
<td>Morbidity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falls</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>N/A(^g)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>6 (5.3)</td>
<td>2 (6.1)</td>
<td>1 (2.2)</td>
<td>3 (8.3)</td>
<td>.46</td>
</tr>
<tr>
<td>Deep vein thrombosis</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\(^a\)Overall cohort refers to the study population from phases 2 to 4. Study demographics are not shown for phase 1 patients.
\(^b\)Phase 2 refers to the period during CPIP implementation (May 2019 to September 2019).
\(^c\)Phase 3 refers to the sustainability phase post-CPIP but before the COVID-19 pandemic (October 2019 to March 2020).
\(^d\)Phase 4 refers to the sustainability phase post-CPIP during the COVID-19 pandemic (April 2020 to September 2020).
\(^e\)POD: postoperative day.
\(^f\)ICU: intensive care unit.
\(^g\)N/A: not applicable.

### Reasons for the Failure of Early Postoperative Mobilization

Table 2 summarizes the reasons for the failure of early postoperative mobilization. A total of 21 patients (18.4%) failed to ambulate \(\geq 30\) m on POD 2. Among these patients, 15 (71.4%) underwent open surgery, and 17 (81.0%) had the ON-Q PainBuster inserted intraoperatively. A total of 13 patients (61.9%) had either inpatient or outpatient prehabilitation before the surgery. The most common reason for failure to achieve the ambulation target was pain (6/21, 28.6%), followed by lethargy or giddiness (5/25, 23.8%). In addition, 2 patients (9.5%) were required to have complete rest in bed due to chest tube removal, and 2 patients (9.5%) had ongoing blood transfusions upon review by the physiotherapist and hence did not ambulate. Another 2 patients (9.5%) were admitted in the intensive care unit and were not stable enough for physiotherapy.

### Discussion

#### Principal Findings

Our study demonstrated the long-term sustainability of the CPIP to promote early mobilization following elective major HPB surgery after CPIP implementation. The mobilization rate during the COVID-19 pandemic was 79%.

CPIPs target a specific, measurable goal, identify critical barriers, and develop a model for improvement. We previously described the success of our CPIP in improving postoperative mobilization [9]. A quality dashboard inclusive of a Pareto chart was provided to clinician stakeholders in 2018. Engagement of physiotherapy and nursing colleagues was done to understand the micro and macro workflows relevant to mobilization. Root cause analysis for barriers to mobilization was done by a core team trained in CPIP use. The surgeon-led multidisciplinary quality improvement initiative with multiple PDSA cycles adhering to the CPIP philosophy led to improved process.
outcomes along with cost savings [9]. However, a CPIP can only be successful if it is sustainable. Sustainability is defined as the capacity of a health service to deliver health care over time with considerations for future generations [10]. It is an essential facet of health care innovation. It has, therefore, been incorporated to be included in the Institute of Medicine’s six domains of quality by the Royal College of Physicians [10]. Alexander et al [11] concluded that the median follow-up time for health care quality improvement projects was less than 1 year, which is insufficient to observe the long-term effects of any implementation on clinical outcomes. We continued the follow-up of our CPIP for 1 year following the implementation and divided our analysis into phase 3 (before COVID-19) and phase 4 (during COVID-19) to observe differences during these two periods. During both periods, we achieved a ≥75% target mobilization rate.

Early mobilization is a facet of the ERAS program; a meta-analysis by Ji et al [20] on the use of ERAS in pancreatic surgery demonstrated lower incidence of delayed gastric emptying (odds ratio (OR) 0.58, 95% CI 0.48-0.72, P < .001), lower postoperative complication rates (OR 0.57, 95% CI 0.45-0.72, P < .001), and shorter length of hospital stay (weighted mean difference -4.45, 95% CI -5.99 to -2.91, P < .001). However, compliance was not reported in the meta-analysis. The failure of ERAS programs may be due to a lack of compliance rather than the concept of ERAS itself [21,22]. Elias et al [23] published the Reporting on ERAS Compliance, Outcomes, and Elements Research (RECOvER) checklist to improve compliance, including the need to describe a strategy for early mobilization. They defined early mobilization as fulfilling all of the following: (1) in the postoperative anesthesia unit, to ambulate from bed to chair, (2) on POD 0, to ambulate 3 times and sit out of bed for all meals (no distance or time duration specified), and (3) on POD 1, to sit out of bed for ≥8 hours. This provides a standardized checklist with a clear definition, but the definition of early mobilization is heterogeneous in other studies. Wind et al [24] defined early mobilization as sitting out of bed for ≥2 hours on POD 0, ≥6 hours on POD 1, and ≥8 hours on POD 2. Gatt et al [25] defined it as sitting out of bed on POD 0 and ambulating the length of the ward on POD 1. We defined early mobilization as sitting out of bed for ≥6 hours on POD 1 and ambulating ≥30 m on POD 2. A review of existing literature on ERAS programs showed heterogeneous definitions of early postoperative mobilization, ranging from “time spent out of bed” and “time ambulated” to “distance or steps walked on POD 2 or beyond.” Hence, the value of “30 m” was chosen based on past experience and the practical needs of patients in our local context: 30 m is the approximate distance to ambulate from the living room to the toilet and back. The ability to do so would suggest that the patient is able to independently carry out activities of daily living, making this a meaningful distance target. Furthermore, the to-and-fro distance from patients’ cubicle to the ward entrance is approximately 30 m, making it logistically easier for physiotherapists and nurses to record the distance ambulated [9,26]. It is important to note that the terminologies “mobilization” and “ambulation” are not synonymous. Patients were required to fulfill both criteria—sitting out of bed for ≥6 hours on POD 1 and ambulating ≥30 m on POD 2—to be deemed successful in early postoperative mobilization. While we agree that patients are instructed to sit out of bed usually on either POD 0 or POD 1, it is the act of ambulating that is more relevant to patients’ physiologic function and activities of daily living. Therefore, we defined mobilization by the act of ambulating, rather than by only sitting out of bed. Further prospective studies examining postoperative mobilization should use standardized and concise definitions of mobilization to have a clear endpoint and for results to be reproducible for large-volume meta-analyses.

The COVID-19 pandemic has resulted in disruption in the delivery of health care services, especially in the surgical subspecialties. Recommendations were made for postponing elective surgical cases where possible [27]. Locally, there was a shift toward nonoperative management for stable, benign conditions such as uncomplicated acute cholecystitis [28]. This was to redirect resources toward the management of patients with COVID-19. Our study, however, showed that we were able to maintain mobilization targets even during the pandemic. Following the CPIP, we continued implementing preoperative counseling and reinforcing the importance of early mobilization on POD 1 during routine ward rounds. Reasons that were previously identified for failure to ambulate continue to be addressed. Pain was the most common reason for failure to achieve the ambulation target (n=6, 28.6%). In our institution, the ON-Q PainBuster was placed intraoperatively in the preperitoneal space for major open surgeries to deliver bupivacaine or ropivacaine through continuous infusion. This is reported to be effective in reducing postoperative pain and facilitating early ambulation compared to a placebo [29]. PCA was also used as part of our multimodal approach for analgesia. Nevertheless, pain remained the most typical reason for failure to achieve ambulation targets; this is likely because of the need to balance the side effects of excessive analgesic use, such as nonsteroidal inflammatory drugs, with the risk of renal impairment adverse cardiac events and gastrointestinal bleeding [30]. Use of opioids is also associated with delayed recovery of bowel function, as well as postoperative nausea and vomiting, which may limit ambulation. Therefore, titration of analgesia needs to obtain the best control of pain and limit side effects to improve mobilization. Interestingly, there was a significant reduction in pain score on POD 2 from 2.3 (SD 1.8) during phase 2 to 1.3 (SD 1.3) during phase 4 (P=.01), with comparable incidence of laparoscopic surgery, which may have contributed to the sustainability of early postoperative mobilization. While improved pain incentivizes patients to mobilize early, Ni et al [31] demonstrated improved pain scores on POD 5 in patients who had early ambulation compared to the control group (mean 3.1, SD 1.1 vs mean 3.8, SD 2.4; P=.05). To add on to the discussion, while epidural analgesia is an alternative for pain control, our institution prefers ON-Q PainBuster to epidural infusion as ON-Q PainBuster is relatively easier to insert and remove and does not require an INR ≤1.4 for safe removal unlike epidural catheter. A systematic review by Munroop et al [32] showed that preperitoneal wound catheters provide statistically, but not clinically, significantly different pain control at rest on POD 1 as epidural analgesia (mean difference 0.44, 95% CI 0.06-0.79; P=.02), with a lower incidence of hypotension (relative risk 0.29, 95% CI 0.13-0.68; P=.004) and patient
satisfaction. While pain is the most common reason for failure of ambulation, we have attempted to mitigate it via adoption of a multipronged pain control approach.

Further, a plausible reason for the sustainability in having high postoperative mobilization rates despite no active oversight could be due to staff empowerment following CPIP implementation. Our CPIP has emphasized the importance of early postoperative mobilization, with the aim of ambulating ≥30 m on POD 2. Chong et al [33], who studied nurses’ practices regarding early mobilization among mechanically ventilated patients, found that the majority of nurses (99.2%) observed in-bed mobilization among patients, but only a minority (14.4%) saw out-of-bed mobilization. They attributed the lack of doctors’ order for physiotherapy or the lack of nursing staff availability as possible reasons for the lack of out-of-bed mobilization [34]. In line with this, we feel that the strong reinforcement of early postoperative mobilization has provided nursing staff with confidence to allow patients to sit out of bed on POD 1 and promote early mobilization where feasible and when permitted by resource availability. Other indirect measures played by nursing colleagues include charting of pain scores and provision of adequate analgesia to manage pain, which is the most common factor for the lack of early postoperative mobilization [35].

Early postoperative mobilization has been shown to improve clinical outcomes, with a reduced length of hospital stay and incidence of pneumonia and deep vein thrombosis [2,3]. Incidence of postoperative pneumonia following liver surgery has been reported to range from 8.2% to 13% [36-39]. Mobilization has been postulated to elicit cardiopulmonary responses, which enhance oxygen transport, increase tidal volume that may reverse atelectasis, and improve gas exchange or reduce the risk of aspiration in view of an upright position [40]. Our study showed a relatively lower 5.3% overall incidence of postoperative pneumonia from phases 2 to 4, which may be multifactorial: laparoscopic surgery, prehabilitation, early mobilization, and multimodal analgesia with adequate pain control [41].

To improve clinical outcomes, it is integral to improve the process outcomes of all integral components of ERAS protocols. Increasing compliance to existing protocols is an important step forward [2]. Our study demonstrated improvement in early postoperative mobilization rates within our institution; however, our sample size is relatively small, and the generalizability of the results is limited due to the heterogeneous patient population. The concept and technology of health information exchange (HIE) may be adopted to improve the situation. HIE as defined as the use of technology to share clinical and administrative data electronically across health care institutions and repositories; it may be considered to facilitate large-scale prospective studies to provide improved quality of health care and cost savings [42]. A novel method of tracking the compliance and development of predictive risk scores for various clinical outcomes was recently developed by Cochran et al [43]. Research Electronic Data Capture (REDCap), which is an electronic data management system primarily used for data collection, was used to track compliance to ERAS protocols in our institute. Use of both health informatics and REDCap simplifies the process of tracking clinical outcomes and disseminating clinical performance indicators. This permits a quick update of the ERAS dashboard, planning of targeted interventions to improve outcomes, and easy sharing of data across institutions through the HIE technology. Furthermore, embracing these technologies reduces missing data and recording bias to some extent. Institutions with ongoing ERAS protocols should also re-examine their respective surgery dashboards to ensure continued quality improvements. Interinstitutional collaboration should also be encouraged to facilitate high-powered evidence.

Strengths and Limitations
The strength of our study is that it is, to our knowledge, the only study to report the long-term sustainability of mobilization during the COVID-19 pandemic. We also included reasons for failure to achieve ambulation targets to improve future CPIPs. However, this study has a few limitations. Heterogeneity of the study population, which includes patients who underwent liver surgery and pancreatic surgery, limits the generalizability of the results. We also did not assess the benefits of early postoperative mobilization on clinical outcomes, such as length of hospitalization and postoperative morbidity [2]. The primary aim of this study was descriptive, to describe our experience in sustaining the CPIP at 1 year following implementation during the COVID-19 pandemic. This study also included patients who underwent prehabilitation, which may indirectly have led to the achievement of the mobilization target.

Conclusion
This follow-up study demonstrated the sustainability of our CPIP in improving early postoperative mobilization rates in patients who underwent elective major HPB surgery 1 year following implementation, even during the COVID-19 pandemic. Further large-scale, multi-institutional prospective studies are needed to define mobilization and assess compliance to early mobilization initiatives. Sustaining a clinical improvement initiative is an essential determinant of value-driven patient-centric health care.

Acknowledgments
We would like to thank all the doctors, nurses, and allied health professionals who were involved in this quality improvement project to improve early postoperative mobilization rates at our institution.

Conflicts of Interest
None declared.
References


Abbreviations

ANOVA: analysis of variance
CPIP: clinical practice improvement program
ERAS: Enhanced Recovery After Surgery
HIE: health information exchange
HPB: hepatopancreatobiliary
INR: international normalized ratio
OR: odds ratio
PCA: patient-controlled analgesia
PDSA: plan-do-study-act
POD: postoperative day
RECOvER: Reporting on ERAS Compliance, Outcomes, and Elements Research
REDCap: Research Electronic Data Capture

©Kai Siang Chan, Bei Wang, Yen Pin Tan, Jaclyn Jie Ling Chow, Ee Ling Ong, Sameer P Junnarkar, Jee Keem Low, Cheong Wei Terence Huey, Vishal G Shelat. Originally published in JMIR Perioperative Medicine (http://periop.jmir.org), 06.10.2021. 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.
Discussion of Weight Loss Surgery in Instagram Posts: Successive Sampling Study

Zoe C Meleo-Erwin¹, MA, PhD; Corey H Basch¹, MPH, EdD; Joseph Fera², PhD; Bonnie Smith¹, BS

¹Department of Public Health, William Paterson University of New Jersey, Wayne, NJ, United States
²Department of Mathematics, Lehman College, City University of New York, Bronx, NY, United States

Corresponding Author:
Zoe C Meleo-Erwin, MA, PhD
Department of Public Health
William Paterson University of New Jersey
300 Pompton Rd
University Hall, Suite 359 – Office 370
Wayne, NJ, 07470
United States
Phone: 1 9737202394
Email: erwinz@wpunj.edu

Abstract

Background: The majority of American adults search for health and illness information on the internet. However, the quality and accuracy of this information are notoriously variable. With the advent of social media, US individuals have increasingly shared their own health and illness experiences, including those related to bariatric surgery, on social media platforms. Previous research has found that peer-to-peer requesting and giving of advice related to bariatric surgery on social media is common, that such advice is often presented in stark terms, and that the advice may not reflect patient standards of care. These previous investigations have helped to map bariatric surgery content on Facebook and YouTube.

Objective: This objective of this study was to document and compare weight loss surgery (WLS)–related content on Instagram in the months leading up to the COVID-19 pandemic and 1 year later.

Methods: We analyzed a total of 300 Instagram posts (50 posts per week for 3 consecutive weeks in late February and early March in both 2020 and 2021) uploaded using the hashtag #wls. Descriptive statistics were reported, and independent 1-tailed chi-square tests were used to determine if a post’s publication year statistically affected its inclusion of a particular type of content.

Results: Overall, advice giving and personal responsibility for outcomes were emphasized by WLS posters on Instagram. However, social support was less emphasized. The safety, challenges, and risks associated with WLS were rarely discussed. The majority of posts did not contain references to facts from reputable medical sources. Posts published in 2021 were more likely to mention stress/hardships of living with WLS (45/150, 30%, vs 29/150, 19.3%; P=.03); however, those published in 2020 more often identified the importance of ongoing support for WLS success (35/150, 23.3%, vs 16/150, 10.7%; P=.004).

Conclusions: Given that bariatric patients have low rates of postoperative follow-up, yet post-operative care and yet support are associated with improved health and weight loss outcomes, and given that health content on the web is of mixed accuracy, bariatric professionals may wish to consider including an online support forum moderated by a professional as a routine part of postoperative care. Doing so may not only improve follow-up rates but may offer providers the opportunity to counter inaccuracies encountered on social media.


KEYWORDS
bariatric surgery; social media; Instagram; health promotion; post-operative medicine; Instagram; online health information; information accuracy; surgery; information quality
Introduction

It is now well understood that health and illness are key drivers of internet use. Pew Research Center investigators have found that nearly three-quarters of American adults go on the internet to look for information related to health and health care [1]. Over one-third of these individuals use the internet in an attempt to self-diagnose [2]. Notably, high rates of web-based health seeking behavior have been found in other economically advanced countries, such as France [3], Germany [4], and Scotland [5]. In the United States, women, White adults, young people, and those of higher socioeconomic status are most likely to seek diagnoses on the web [6]. Pew Research Center indicates that the most popular web-based health searches include treatments for specific health concerns and conditions [1]. It is therefore unsurprising that the internet has become a primary source of information for individuals who desire weight loss in general [7] and by means of bariatric surgery procedure in particular [8,9]. Furthermore, it seems that as weight increases, so does the likelihood of searching on the web for weight loss solutions [10]. Finally, there is some evidence that individuals interested in bariatric surgery use the internet for health information-seeking more than individuals seeking other forms of digestive system surgery [8].

The (relative) anonymity of internet use may help to explain why those medically diagnosed as overweight or obese—defined by the US Centers for Disease Control [11] as having a BMI of “25.0 to <30” or “30.0 or higher,” respectively—are particularly inclined to use the internet for these purposes. Individuals with stigmatized health conditions have been found to be strong users of the internet for health and illness information searches [12]. Weight-related bias, stigma, and discrimination are considered pervasive in such realms as media, education, employment, and health care [13,14]. This inequitable treatment adversely impacts both the life chances and health outcomes of higher-weight individuals [13-16]. The World Health Organization [17] notes that there is a lack multinational studies comparing weight bias in different countries. However, similar rates of weight bias have been reported in the United States, Canada, Iceland, and Australia [18]. Studies conducted in both the United States [19] and Germany [20] show a positive association between body weight and weight-related bias. Specifically, as an individual’s weight increases, so does the degree of weight-based bias that they experience.

In health care, individuals medically classified as having obesity often face a reduced quality of care, in part due to negative provider attitudes and in part because many health care facilities are not well suited to treat larger-bodied people [13]. Concerns over potential provider bias may lead to delays in seeking care for higher-weight individuals [21]. Thus, the ability to learn more on the web about weight loss in general and weight loss surgery (WLS) in particular may seem to some to be a safer option. Of concern, however, is that the quality and accuracy of health information on the web varies depending on the source. Numerous researchers [22,23] have demonstrated variability in web-based WLS information, noting that the risks of bariatric procedures are not well discussed and the content often lacks provider input.

Just as the advent of the internet has radically shifted the modes by which individuals seek health and illness information, the rise of social media has transformed the ways in which individuals express their own experiences and connect with others around these topics. For example, Pew Research Center investigators [1] have found that going on the web to read or watch videos about others’ health and illness experiences, finding people with similar health concerns, asking health and illness-related questions, and posting about one’s own health and illness experiences are common drivers of social media use, especially for those with one or more chronic health issues. Although the majority of Americans participate in social media, platform use varies by age; older US adults tend to frequent Facebook and YouTube, while younger individuals prefer sites such as Snapchat and Instagram [24].

Researchers have begun to document WLS-related content on social media sites. For example, some investigations have described the nature of WLS support forums on Facebook. Kombrahl et al [25] determined that patient posts most commonly pertained to information seeking, provision of tips and advice, and lending encouragement and support to others. In a separate study, Kombrahl et al [26] noted that individuals in Facebook WLS patient support groups most commonly solicited and shared nutrition-related advice. Much of the advice given was presented in stark terms (“eat this,” “don’t eat that”), as opposed to the more nuanced and personalized approach that registered dieticians take when advising patients. The authors of an additional investigation [27] found that although Facebook WLS support group membership includes bariatric providers, these individuals had low levels of participation. Recommendations were instead often provided by seasoned patients who positioned their time out of surgery as grounds for providing advice and information to preoperative and newly postoperative individuals. Bariatric professionals do have a presence on YouTube; however, as with Facebook, content is largely driven by lay individuals, and the most popular videos tend to be of lower educational quality [28].

Although these studies have helped to map the nature of WLS content on social media, there remains a dearth of research on Instagram specifically. Instagram is considered to be one of the most commonly used social media sites in operation, with over 1 billion users worldwide [29]. According to the most recent study on social media usage by the Pew Research Center, Instagram is the third most frequently used social media platform in the United States [30]. Globally, Instagram has been ranked fifth in terms of numbers of active users [31]. Instagram users likely span a variety of demographic categories; however, they tend to be younger individuals [23]. For individuals aged 18 to 29 years, Instagram is the second most frequently used social media app in the United States, after YouTube [30]. Instagram is also the social media platform with the second largest age differential between older and younger American users [30]. One investigation found that Instagram users are also disproportionately female, of lower socioeconomic status, and from urban areas [24].

Beyond the general lack of research on this topic, an examination of WLS-related content on Instagram is warranted for a number of reasons. First, although the average age of US
bariatric patients has been documented as approximately mid-40s [32], the incidence of WLS in young people increased during the early 20th century [33,34]. Second, individuals between the ages of 18 and 30 years have been found to rely on the internet, including social media, for overall health information [35]. There is evidence to suggest that this age group is more likely to do so than older age cohorts [36]. Third, social media platforms are now commonly used by individuals to seek and relate information related to WLS [7-9; 25-27]. Fourth, seekers of web-based health information do not simply passively take in such information on WLS but actively make decisions based upon what they read on the web [9,22]. These individuals continue to use the internet for postoperative WLS support [9]. Finally, given the shift away from the routine provision of in-person care and support necessitated by the COVID-19 pandemic, bariatric preoperative and postoperative patients may have increasingly turned to social media to fill this gap. This study therefore aimed to document and compare WLS-related content on Instagram in the weeks leading up to the COVID-19 pandemic lockdowns to those exactly 1 year later.

Methods

The methods used in this study were derived from prior studies that investigated health content posted on social media [37,38]. The sample was comprised of 300 Instagram posts created by users with the hashtag #wls. This hashtag was chosen because it was included in by far the most posts compared to any hashtag associated with weight loss surgery or bariatric surgery (approximately 1 million more posts). The investigation proceeded by means of a successive sampling study: half of the sample was collected in 2020, and the other half was collected exactly 1 year later in 2021. During both 2020 and 2021, 50 posts per week were collected over the course of 3 weeks (the last week of February, the first week of March, and the second week of March), thus minimizing the chance that multiple posts were originating from the same user. Captions, hashtags, and comments made by the user with the same username associated with the post were included in the coding. Posts with photos and videos were included in this study, and any included text was factored into the coding of content. Exclusion criteria included posts that were in a language other than English, images posted without an explanation, and any posts aimed at advertising or selling a product.

A description and the date posted were noted for each post. Each of the posts was studied to see whether it exhibited a predetermined content category. Content coding categories were based on findings of prior research on internet forum and social media use by bariatric surgery patients [25,26,39]. For the data collected, descriptive statistics were recorded, and independent 1-tailed chi squared tests ($\alpha=0.05$) were performed to determine if the year of the post statistically impacted the presence of a given characteristic in the post. Data entry, organization, and analysis were performed in Excel (Microsoft Corporation). The Institutional Review Board at William Paterson University does not require review for studies that involve publicly available social media content. Nevertheless, users may have an expectation of privacy when posting on Instagram; thus, no usernames have been reported in this study.

Results

Table 1 includes a list of 15 different content characteristics of the studied posts and indicates how many of the 300 posts included this content. The table also includes a breakdown of these counts by post year, with relative percentages indicated.
The majority of the posts included a personal story (275/300, 91.7%) and/or included a photo of a person (272/300, 90.7%). Exactly 40% of the posts (120/300) identified the type of WLS, while just under 40% (115/300, 38.3%) gave advice, tips, or suggestions. Just under one-third of the posts (95/300, 31.7%) stated body image was a reason the user underwent WLS, and just over 25% of the posts (n=5, 26.3%) indicated health/pregnancy as a reason. In 2021, 19 of the 150 posts (12.7%) included a reason for the WLS. Of these 19 posts, 10 (52.6%) stated body image was a reason the user underwent WLS, and 5 (38.5%) indicated health/pregnancy as a reason. In 2020, 13 of the 150 posts (8.7%) included a reason for the WLS. Of these 13 posts, 6 (46.2%) stated body image as a reason the user underwent WLS, and 5 (38.5%) indicated health/pregnancy as the reason for undergoing WLS.

Table 1. Observed characteristics and content of 300 Instagram posts on WLS (150 from 2020 and 150 from 2021).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (N=300), n (%)</th>
<th>2020 (n=150), n (%)</th>
<th>2021 (n=150), n (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal story</td>
<td>275 (91.7)</td>
<td>140 (93.3)</td>
<td>135 (90)</td>
<td>.30</td>
</tr>
<tr>
<td>Photo of person</td>
<td>272 (90.7)</td>
<td>127 (84.7)</td>
<td>145 (96.7)</td>
<td>&lt;.001^a</td>
</tr>
<tr>
<td>Identifies type of WLSb</td>
<td>120 (40)</td>
<td>95 (63.3)</td>
<td>25 (16.7)</td>
<td>&lt;.001^a</td>
</tr>
<tr>
<td>Gives advice, tips, suggestions</td>
<td>115 (38.3)</td>
<td>63 (42)</td>
<td>52 (34.7)</td>
<td>.19</td>
</tr>
<tr>
<td>Stresses personal responsibility for improving health</td>
<td>95 (31.7)</td>
<td>66 (44)</td>
<td>29 (19.3)</td>
<td>&lt;.001^a</td>
</tr>
<tr>
<td>Mentions stress/hardships of living with WLS</td>
<td>74 (24.7)</td>
<td>29 (19.3)</td>
<td>45 (30)</td>
<td>.03^c</td>
</tr>
<tr>
<td>Identifies number of years out of surgery</td>
<td>69 (23)</td>
<td>36 (24)</td>
<td>33 (22)</td>
<td>.26</td>
</tr>
<tr>
<td>Identifies importance of ongoing support for WLS success</td>
<td>51 (17)</td>
<td>35 (23.3)</td>
<td>16 (10.7)</td>
<td>.04^d</td>
</tr>
<tr>
<td>Identifies reason for WLS</td>
<td>32 (10.7)</td>
<td>13 (8.7)</td>
<td>19 (12.7)</td>
<td>.26</td>
</tr>
<tr>
<td>Identifies ongoing side effects of surgery</td>
<td>25 (8.3)</td>
<td>17 (11.3)</td>
<td>8 (5.3)</td>
<td>.06</td>
</tr>
<tr>
<td>Identifies postoperative complications from surgery</td>
<td>23 (7.7)</td>
<td>17 (11.3)</td>
<td>6 (4)</td>
<td>.02^c</td>
</tr>
<tr>
<td>Identifies ongoing work of WLS/WLS as “just a tool”</td>
<td>19 (6.3)</td>
<td>15 (10)</td>
<td>4 (2.7)</td>
<td>.009^d</td>
</tr>
<tr>
<td>States facts</td>
<td>7 (2.3)</td>
<td>1 (0.7)</td>
<td>6 (4)</td>
<td>.06</td>
</tr>
<tr>
<td>Discusses weight regain</td>
<td>4 (1.3)</td>
<td>1 (0.7)</td>
<td>3 (2)</td>
<td>.31</td>
</tr>
<tr>
<td>Addresses anti-WLS stigma</td>
<td>3 (1)</td>
<td>2 (1.3)</td>
<td>1 (0.7)</td>
<td>.56</td>
</tr>
</tbody>
</table>

^aSignificant at P<.001.
^bWLS: weight loss surgery.
^cSignificant at P<.05.
^dSignificant at P<.01.

Independent 1-tailed chi-square tests (α=.05) were performed to determine if a post’s publication year (2020 vs 2021) statistically affected its inclusion of a particular content characteristic. The last column of Table 1 provides the resulting P values from these tests, with footnotes indicating statistically significant results with P<.05, P<.01, or P<.001. Compared to 2021, posts made in 2020 more frequently stressed personal responsibility for improving health (55/150, 44%, vs 29/150, 19.3%; P<.001), identified the type of WLS (95/150, 63.3%, vs 25/150, 16.7%; P<.001), identified the importance of ongoing support for WLS success (23.3% vs 10.7%; P=.004), identified postoperative complications from surgery (17/150, 11.3%, vs 6/150, 4%; P=.02), and identified ongoing work of WLS/WLS as “just a tool” (15/150, 10%, vs 4/150, 2.7%; P=.009). Posts published in 2021, however, were more likely to include the following content as compared to posts made in 2020: a photo of a person (145/150, 96.7%, vs 127/150, 84.7%; P<.001) and mention of stress/hardships of living with WLS (45/150, 30%, vs 29/150, 19.3%; P=.03).

Of the 150 posts from 2020, 13 (8.7%) included a reason for the WLS. Of these 13 posts, 6 (46.2%) stated body image as a reason that the user underwent WLS, and 5 (38.5%) indicated health/pregnancy as a reason. In 2021, 19 of the 150 posts (12.7%) included a reason for the WLS. Of these 19 posts, 10 (52.6%) stated body image was a reason the user underwent WLS, and just over 25% of the posts (n=5, 26.3%) indicated health/pregnancy as the reason for undergoing WLS.

**Discussion**

This descriptive, successive sampling study documented and compared Instagram content on bariatric surgery in the weeks leading up to the COVID-19 pandemic lockdown to those posted during these same weeks 1 year later. In both years, the majority of individuals who had undergone bariatric surgery did not note their reason for doing so. However, when those reasons were mentioned, they more commonly pertained to body image issues and concerns over future health risks than to current weight-related health problems. Posts containing facts—wherein a user referred to an external medical source of the information contained in the post—were very rare. By contrast, posts offering personal advice, tips, and suggestions for others were far more common. This is understandable, however, given that previous research [25] has found that giving of WLS advice on social media forums is often presented in a more stark and less nuanced

https://periop.jmir.org/2021/2/e29390
manner than a bariatric professional might otherwise provide, it is possible that the uptake of personal advice offered on social media will have negative consequences for bariatric patients. This may be especially the case for younger individuals who have lower levels of health literacy [40,41].

Interestingly, although references to the risks and challenges (eg, complications, side effects, weight regain) associated with WLS were infrequent, discussion of the safety profile of WLS procedures did not occur at all. Personal responsibility for health and weight loss outcomes was stressed more commonly than the importance of social support. Finally, despite research documenting the extentiveness of weight bias, stigma, and discrimination [13,20], as well as research which has found that such stigma extends to an individual’s decision to have bariatric surgery (with WLS patients being accused of having “taken the easy way out” by choosing a surgical means of weight loss) [42,43], psychosocial factors related to either the decision to have WLS or living with such a procedure were uncommon.

Comparing the two years, posters in 2021 were more likely to emphasize the ongoing challenges of living with a WLS procedure but less likely to emphasize personal responsibility for health or the ongoing work required to live with a WLS procedure (to avoid complications, side effects, or weight regain). Moreover, in 2021, posts highlighting the importance of support for postoperative success decreased compared to the year prior. Taken together, it appears that in 2021, bariatric patients posted more often about challenges they faced regarding WLS but less often about intrapersonal or clinical resources that might help address these challenges. This finding is notable given the timing of the data collection. As noted, the first set of data were collected in the weeks immediately preceding COVID-19–related lockdowns in the United States and worldwide. By the time the data were collected 1 year later in 2021, over 500,000 Americans had died of the disease [44]. Globally, there were nearly 2.5 million cumulative deaths from COVID-19 [45]. Additionally, during this time, highly contagious COVID-19 variants were circulating; however, mass vaccination efforts were beginning to move forward. Some preliminary evidence has found that bariatric patient follow-up rates may have been improved by a shift to telemedicine during the pandemic [46]. However, other studies suggest that the pandemic may have increased the risk of adverse physical and mental health outcomes for bariatric patients [47-50]. Clearly, more research will be needed to determine the long-term impact of the pandemic on bariatric patient outcomes.

Arguably, the impact of using the internet to seek health and illness information and connect with others who share concerns and experiences about health and illness is mixed, both in general and regarding WLS specifically. The internet has democratized access to information regarding health and illness (as well as other topics). This democratization has likely enhanced health-related empowerment and self-efficacy and has facilitated patient advocacy, particularly around contested health issues [51]. Additionally, social media forums offer individuals a convenient way to give and receive ongoing support. Given that logistical barriers (eg, availability, time, location, competing responsibilities, and associated expenses) may prevent some individuals from accessing in-person support groups, online support forums can help meet this vital need. In this vein, and given that ongoing support has been found to be associated with improved health outcomes postoperatively [52], we view the use of Instagram and other social media by individuals who undergo WLS to have potential positive benefits.

However, as noted, health and illness information on the web, including that related to bariatric procedures specifically, is of varied quality and accuracy. There does seem to be some debate in the literature regarding whether individuals searching for such information find health information on the web to be credible compared to that from professional sources. What is well known, however, is that the majority of individuals who undergo WLS in the United States do not return for follow-up care and support services within the first 1 to 2 years after surgery [52,53]. Patients may therefore rely on peer-generated information and advice without verifying the accuracy of that information or discussing the appropriateness of applying it to themselves with their bariatric providers.

It is in this the context that we situate the findings of this study and our concern therein that Instagram posters emphasized personal responsibility but not ongoing support for health and weight loss outcomes. More particularly, we are concerned that patients may experience challenges as well as more serious side effects and complications of undergoing WLS but may rely on the application of peer-to-peer advice and information that does not meet standards of care to address these concerns. This may be especially the case for young adults, who have been shown to adopt health and eating behaviors based on Instagram content even when they understand that the posted images and content on the site are highly curated [41]. For our purposes, such curated images may include representations of visual transformation posted without any accompanying fact-based discussion. Young people may also be particularly vulnerable to health advice coming from “influencers” or aspiring influencers whose key purpose for posting social media content is to gain followers and paid sponsorships [54]. Although our study did not assess whether the posters were influencers, given that all posts included in this study were publicly available, it is possible that some of the individuals were posting in this vein.

A number of limitations of this study should be noted. First, in our investigation, we collected data at two moments in time: late February to mid-March 2020, and exactly 1 year later in 2021. Given the constantly changing nature of social media, it is possible that the content would be different at another point in time. Second, this study relied on publicly available posts. It could therefore not capture the nature of posts made by private accounts. However, as noted, our study may have been more likely to capture posts made by influencers and aspiring influencers. Third, we investigated the content of WLS Instagram posts but not the comments. Therefore, we cannot make claims about the conversations that occur between individuals on Instagram regarding WLS. Fourth, we cannot state what impacts, if any, the posts included in the sample had on Instagram users who engaged with them. Fifth, although Instagram is a leading social media site, our study only investigated WLS content on this one platform rather than across social media sites. Sixth, our cutoff of 50 posts at each data...
from credible medical sources. Although the safety profile of
WLS procedures has improved over time, complications and
side effects still occur with some frequency [55]. Moreover,
although side effects and complications of WLS are not
reducible to patient behavior, they are unquestionably related.
Given that American bariatric patients have low rates of
postoperative follow-up [52,53], and yet ongoing post-operative
support is linked with improved health outcomes [52], offering
weight loss surgery patients a variety of convenient and
accessible options for accessing care and support is warranted.
In 2011, Kaiser et al [56] argued that offering a mix of formats
for WLS clinical support groups—telephone, internet, and
in-person—may help address barriers to attending in-person
group sessions. Offering patients an active online support forum
moderated by a bariatric professional may additionally offer
the opportunity to counter misinformation that circulates in
peer-led forums and may also provide the sort of balanced and
nuanced information that can help patients navigate any ongoing
challenges of living with a bariatric surgical procedure.

Acknowledgments
We are grateful to the anonymous reviewers who provided helpful comments on an earlier version of this manuscript. The manuscript is much improved thanks to their efforts. No funding was received for this investigation.

Conflicts of Interest
None declared.

References
11. US Centers for Disease Control and Prevention. URL: [https://www.cdc.gov/obesity/adult/defining.html], [accessed 2021-06-02]


29. Instagram. URL: https://about.instagram.com/about-us [accessed 2021-04-05]


**Abbreviations**

WLS: weight loss surgery
Patient Engagement in the Design of a Mobile Health App That Supports Enhanced Recovery Protocols for Cardiac Surgery: Development Study

Anna M Chudyk1, BHSc, MSc, PhD; Sandra Ragheb2, BN; David Kent3, BKin, MSc; Todd A Duhamel3,4, PhD; Carole Hyra5; Mudra G Dave3; Rakesh C Arora6*, BA, MD, PhD, FRCSC, FACS; Annette SH Schultz7,8*, RN, PhD

1Department of Family Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada
2Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada
3Faculty of Kinesiology and Recreation Management, University of Manitoba, Winnipeg, MB, Canada
4Institute of Cardiovascular Sciences, St Boniface Hospital, Winnipeg, MB, Canada
5Healthy Heart Patient and Caregiver Researcher Group, Institute of Cardiovascular Sciences, St Boniface Hospital, Winnipeg, MB, Canada
6Department of Surgery, Section of Cardiac Surgery, Max Rady College of Medicine, University of Manitoba, Winnipeg, MB, Canada
7College of Nursing, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada
8Health Services & Structural Determinants of Health Research, St Boniface Research Centre, Winnipeg, MB, Canada
*these authors contributed equally

Corresponding Author:
Anna M Chudyk, BHSc, MSc, PhD
Department of Family Medicine
Rady Faculty of Health Sciences
University of Manitoba
CR3023-369 Tache Avenue
Winnipeg, MB, R2H 2A6
Canada
Phone: 1 7783876969
Email: anna.chudyk@umanitoba.ca

Abstract

Background: Despite the importance of their perspectives, end users (eg, patients, caregivers) are not typically engaged by academic researchers in the development of mobile health (mHealth) apps for perioperative cardiac surgery settings.

Objective: The aim of this study was to describe a process for and the impact of patient engagement in the development of an mHealth app that supports patient and caregiver involvement with enhanced recovery protocols during the perioperative period of cardiac surgery.

Methods: Engagement occurred at the level of consultation and took the form of an advisory panel. Patients who underwent cardiac surgery (2017-2018) at St. Boniface Hospital (Winnipeg, Manitoba) and their caregivers were approached for participation. A qualitative exploration determined the impact of patient engagement on the development (ie, design and content) of the mHealth app. This included a description of (1) the key messages generated by the advisory panel, (2) how key messages were incorporated into the development of the mHealth app, and (3) feedback from the developers of the mHealth app about the key messages generated by the advisory panel.

Results: The advisory panel (N=10) generated 23 key messages to guide the development of the mHealth app. Key design-specific messages (n=7) centered around access, tracking, synchronization, and reminders. Key content-specific messages (n=16) centered around medical terms, professional roles, cardiac surgery procedures and recovery, educational videos, travel, nutrition, medications, resources, and physical activity. This information was directly incorporated into the design of the mHealth app as long as it was supported by the existing functionalities of the underlying platform. For example, the platform did not support the scheduling of reminders by users, identifying drug interactions, or synchronizing with other devices. The developers of the mHealth app noted that key messages resulted in the integration of a vast range and volume of information and resources instead of ones primarily focused on surgical information, content geared toward expectations management, and an expanded focus to include caregivers...
and other family members, so that these stakeholders may be directly included in the provision of information, allowing them to be better informed, prepare along with the patient, and be involved in recovery planning.

Conclusions: Patient engagement may facilitate the development of a detail-oriented and patient-centered mHealth app whose design and content are driven by the lived experiences of end users.

(JMIR Perioper Med 2021;4(2):e26597) doi:10.2196/26597

KEYWORDS
cardiac surgery; perioperative care; enhanced recovery protocols; mobile app; smartphone app; mHealth; development; patient and public involvement; patient engagement in research

Introduction

Enhanced recovery protocols (ERPs) are evidence-based care pathways aimed at standardizing perioperative care. In offering a multimodal and interdisciplinary approach to care, these protocols have been proposed as a clinical strategy to effectively address complex and multisystem vulnerabilities [1,2], like those commonly present in older adults undergoing cardiac surgery [3,4]. Mobile health (mHealth) refers to medical and public health practice supported by mobile devices (eg, smartphones, tablets, patient monitoring devices) [5]. mHealth apps have the potential to enhance the utility of ERPs by increasing the effectiveness of information delivery and patients’ (and caregivers’) retention of information regarding their health care plan [6,7]. There is some evidence to support the feasibility of using mHealth during inpatient recovery of patients who had undergone cardiac surgery [8]. However, researchers’ efforts to develop mHealth for the perioperative cardiac surgery setting (and in general) are often limited by the lack of involvement of end users (such as patients and caregivers) in research activities [9].

Patients, caregivers, and other health service users may be involved in mHealth development studies as research participants or coresearchers, using participatory methods such as user-centered design, the participatory action research framework, and the Center for eHealth Research and Disease Management Roadmap [10]. Patient engagement (also commonly referred to as patient and public involvement, patient involvement, and stakeholder engagement) in research is a form of participatory action research that involves the “coproduction” of research with patients and caregivers. It has been defined as the formation of meaningful and active collaborations between researchers and patients (including informal caregivers) in research governance, priority setting, conduct, and knowledge translation [11]. Lack of attention to end users’ perspectives during the development phase is one of the competing explanations for the relatively low uptake of mHealth by patients [9]. Thus, an important step toward more widespread adoption of patients and caregivers as coproducers of mHealth research is one that facilitates a better understanding of processes for engaging patients and caregivers in mHealth development studies.

This study was set within the context of a Canadian clinical research hospital where our research group is involved in the development and implementation of ERPs for cardiac surgery. As part of this work, we initiated a project that developed an mHealth app and determined its effectiveness in improving knowledge delivery of patient education materials and patient adherence to ERPs during the perioperative period of cardiac surgery. A feasibility study of the mHealth app is currently under review. This study focuses on the patient engagement process employed to develop the mHealth app, which was guided by the Canadian Institutes of Health Research’s Patient Engagement Framework [11] and our scoping review of models and frameworks of patient engagement in health services research [12]. Given the novelty of engaging patients as coproducers of mHealth in academic research settings and among most of our team members, this study aimed to describe a process for and the impact of patient engagement on the development of an mHealth app that supports ERPs for cardiac surgery.

Methods

Ethical Approval and Consent

This study was set in an academic tertiary care center that performs cardiac surgery (St. Boniface Hospital, Winnipeg, Manitoba). Ethical approval for this study was obtained from the University of Manitoba Research Ethics Board as well as the Research Review Committee at St. Boniface Hospital. Patients and caregivers provided written informed consent and were compensated CAD $50 (CAD $1=US $0.80, for time and transportation) in addition to the cost of parking per meeting that they attended. The Guidance for Reporting Involvement of Patients and the Public long-form checklist guided the reporting of patient engagement in this paper [13].

Overview of the mHealth App

The mHealth under development was an app-based platform hosted by BeeWell Health [14]. This study gathered, adapted, and electronically formatted patient-and-caregiver derived content that addressed the patient journey from initial cardiac surgery consent through to the 8-week postoperative recovery period for delivery via the mHealth app. This content targeted 3 aspects of perioperative care (ie, patient-tailored education, optimization of patient health, and patient engagement in care) and focused on 4 domains of information (ie, nutrition, medications, resources, and physical activity). The 4 domains of information targeted by the mHealth app were informed by our previous work with patients who had undergone cardiac surgery and their caregivers (data unpublished). Specifically, focus group sessions identified these areas as priorities for patients who had undergone cardiac surgery and their caregivers. Continued research (ie, web-based and telephone surveys) validated these findings within a larger patient and caregiver
Description of the Patient Engagement Process

Patient engagement in research encompasses a wide range of activities and participation types, as influenced by the characteristics of a given project (eg, scope, time, financial resources) and the contributions patients are willing to offer [11,15-17]. In this study, engagement took the form of an advisory panel and occurred at the level of consultation [17]. The role of the advisory panel was to inform the development (operationalized as design and content) of the mHealth app. The advisory panel met in-person 3 times, approximately 2 weeks apart. Each meeting was approximately 3 hours in duration. Figure 2 displays an outline of the activities that occurred at each meeting. The activities that occurred within the meetings were not only developed to gather advisory panel input on the design and content of the mHealth app but also to create/facilitate an environment that supported the guiding principles that underlie patient engagement (ie, mutual respect, inclusiveness, cobuilding, support; see Multimedia Appendix 1 [11,18-20] for information on our approach to creating an environment that embodied these guiding principles). The primary method used to obtain advisory panel members’ input was group discussions. These discussions centered around 2 open-ended questions: “what information stuck out as important during your patient journey” and “what information do you wish you had known during your patient journey.” In addition, the scope of the discussions was narrowed to 4 domains of information (ie, nutrition, medications, resources, and physical activity) identified through previous work, as well as to the content and layout of information presented in a downloadable generic version of the mHealth app. A skilled facilitator (DEK) led the meetings based on a developed facilitation guide. A notetaker (MGD) and an audio recorder documented the meeting proceedings.

Recruitment

Unlike study participants, patients’ and caregivers’ role in patient engagement activities was to represent lived experiences rather than be representative of them [21]. Thus, given the focus of the mHealth app under development, advisory panel membership was based upon the shared experience of having undergone or cared for someone who had undergone cardiac surgery at our study hospital. Specifically, patients who underwent the cardiac surgery procedure within the previous 2 years (2017-2018) at the study hospital and consented to be listed in a database of individuals interested in participating in future research and their caregivers were approached for advisory panel membership. As women are underrepresented in cardiac research and to obtain perspectives that spanned the gamut of cardiac surgery procedures most typically carried out...
at our study hospital, panel members were selectively chosen for diversity in sex and procedure type. Individuals were excluded if they could not read or communicate in English. Recruitment was targeted at 10-12 individuals based on our and others’ experiences with group dynamics and group size. For example, advisory panels within Patient-Centered Outcomes Research Institute range between 10 and 24 members [22], whereas group sizes of 9-12 and 6-12 are commonly recommended for group processes focused on idea generation and discussion, such as the nominal group technique [23] and focus groups [24], respectively. Smaller group sizes (n=4-12) are large enough to facilitate discussion while leaving room for balanced participation [25].

Impact of Patient Engagement on mHealth App Development

A qualitative exploration was undertaken to determine the impact of patient engagement on the development of the mHealth app. This included description of (1) the key messages generated by the advisory panel, (2) how key messages were incorporated into the development of the mHealth app, and (3) feedback from the developers of the mHealth app about the key messages generated by the advisory panel.

Analysis

Discussions that occur as part of patient engagement activities do not typically produce data that are thematically analyzed [25], as the purpose of patient engagement is to learn from patient experiences and not interpret patient experiences through the researcher’s lens. Thus, “real-time processing” of information takes place during discussions, and the information that is gathered is generally presented as a list of stakeholder-made recommendations used to support project decision making [25]. Accordingly, the meeting facilitator (DEK) employed common techniques (eg, summarization, reflection, asking clarifying questions) to identify advisory panel members’ key messages during discussions. Two study team members (DEK and AMC) reviewed the research assistant’s notes from all 3 meetings along with transcripts from the second meeting to generate a list of key messages about the design and content of the mHealth app. These key messages were presented by a study team member (DEK) to the developers of the mHealth platform to guide the design and content of the mHealth app. In addition, advisory panel members’ sociodemographic characteristics, as obtained from our database of individuals interested in participating in future research (patients) and self-report (caregivers), were summarized with medians (25th and 75th percentiles) or counts (percentages). These descriptive statistics were calculated using Stata version 13.0 (Stata Corp).

Results

Sociodemographic Characteristics of the Participants

Ten individuals (6 patients and 4 caregivers) participated in the advisory panel. The select sociodemographic characteristics of the advisory panel members are shown in Table 1. Each caregiver (n=4) was a patient’s (n=4) spouse. Two of the patients did not have a caregiver attend any of the advisory panel sessions.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients (n=6)</th>
<th>Caregivers (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), median (IQR)</td>
<td>74 (72-76)</td>
<td>N/A</td>
</tr>
<tr>
<td>Females, n (%)</td>
<td>3 (50)</td>
<td>3 (75)</td>
</tr>
<tr>
<td>Ethnicty, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/Caucasian/European</td>
<td>5 (83)</td>
<td>4 (100)</td>
</tr>
<tr>
<td>First Nations/Inuit/Metis</td>
<td>1 (17)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Procedure type, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aortic valve replacement</td>
<td>3 (50)</td>
<td>N/A</td>
</tr>
<tr>
<td>Aortic valve replacement/coronary artery bypass grafting</td>
<td>1 (17)</td>
<td>N/A</td>
</tr>
<tr>
<td>Aortic valve replacement/mitral valve replacement</td>
<td>1 (17)</td>
<td>N/A</td>
</tr>
<tr>
<td>Mitral valve replacement</td>
<td>1 (17)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*N/A: not applicable.

Key Messages About the Design of the mHealth App

A summary of the advisory panel members’ key messages about the design and content of the mHealth app is shown in Table 2. Key messages were about the design features of the mHealth app related to access, tracking, synchronization, and reminders. Specific key messages about the mHealth app design are shown in Table 2.

Table 1. Select sociodemographic characteristics of the advisory panel members (N=10).

Table 2. Key messages about the design of the mHealth app.
Table 2. Key messages about the design of the mobile health app.

<table>
<thead>
<tr>
<th>Key messages (the design of the app should include the ability to…)</th>
<th>Overarching message category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access information ahead of medical appointments</td>
<td>Access</td>
</tr>
<tr>
<td>Access information offline</td>
<td>Access</td>
</tr>
<tr>
<td>Share access to the mobile health app with caregivers and family</td>
<td>Access</td>
</tr>
<tr>
<td>Track prescribed medications and exercises that are assigned both in hospital and during outpatient rehabilitation</td>
<td>Tracking</td>
</tr>
<tr>
<td>Synchronize information from medical devices</td>
<td>Synchronization</td>
</tr>
<tr>
<td>Schedule reminders to take medications</td>
<td>Reminders</td>
</tr>
<tr>
<td>Provide daily reminders about assigned exercises and general physical activity recommendations</td>
<td>Reminders</td>
</tr>
</tbody>
</table>

**Key Messages About the Content of the mHealth App**

During discussions of the study’s 2 open-ended questions and the generic version of the mHealth app, content-specific messages centered around medical terms, professional roles, information specific to cardiac surgery procedures and recovery, educational videos, and travel before/after surgery. When discussing the study’s predefined categories of information, key content-specific messages about (1) nutrition related to what to eat, (2) medications, including drug interactions, (3) resources, including medical devices, and (4) physical activity related to addressing fears, as well as providing information, recommendations, and instructions were generated by the advisory panel. Specific key messages about the mHealth app content are shown in Table 3.

Table 3. Key messages about the content of the mobile health app.

<table>
<thead>
<tr>
<th>Key messages (the app’s content should include…)</th>
<th>Overarching message category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitions of key terms</td>
<td>Medical terms</td>
</tr>
<tr>
<td>Cardiac surgery team contact information</td>
<td>Professional roles</td>
</tr>
<tr>
<td>Information about the functions of the different operating room personnel</td>
<td>Professional roles</td>
</tr>
<tr>
<td>Information specific to the different cardiac surgery procedures</td>
<td>Cardiac surgery procedures</td>
</tr>
<tr>
<td>Information about postoperative recovery, including why you might have a chest tube</td>
<td>Cardiac surgery recovery</td>
</tr>
<tr>
<td>Videos that explain the different cardiac surgery procedures</td>
<td>Educational videos</td>
</tr>
<tr>
<td>Information about driving/traveling after cardiac surgery</td>
<td>Travel</td>
</tr>
<tr>
<td>Instructions on what to eat during the perioperative period</td>
<td>Nutrition</td>
</tr>
<tr>
<td>Recipes geared toward those who are looking to adopt a more heart-healthy lifestyle</td>
<td>Nutrition</td>
</tr>
<tr>
<td>Potential drug interactions</td>
<td>Medications</td>
</tr>
<tr>
<td>Resources for medical devices</td>
<td>Resources</td>
</tr>
<tr>
<td>Information that helps address fears around engaging in physical activity before and after cardiac surgery</td>
<td>Physical activity</td>
</tr>
<tr>
<td>Information about and instructions on the types of physical activities patients can and cannot engage in (specific to procedure and perioperative period)</td>
<td>Physical activity</td>
</tr>
<tr>
<td>Instructions on the physical activity and specific exercises a patient should do if they miss a cardiac rehabilitation session</td>
<td>Physical activity</td>
</tr>
<tr>
<td>Instructions on how to complete exercises assigned both in hospital and during outpatient rehabilitation</td>
<td>Physical activity</td>
</tr>
<tr>
<td>General physical activity recommendations</td>
<td>Physical activity</td>
</tr>
</tbody>
</table>

**Incorporation of the Key Messages Into the Development of the mHealth App**

Key messages about the design and content of the mHealth app were compiled and sent to the mHealth app developers by the study coordinator (DEK). These were then directly incorporated into the mHealth app as long as they could be supported by the existing functionalities of the underlying platform. For example, the platform did not support the scheduling of reminders by users, identifying drug interactions, or synchronizing with other devices. Verbal and written feedback from the mHealth app developers indicated that the key messages were a richer source of information and provided more guidance than typically received from past clients. In particular, the mHealth app developers noted that key messages resulted in the integration of a vast range and volume of information and resources, instead of ones primarily focused on surgical information, content geared toward expectations management, and an expanded focus of the mHealth app to include caregivers and other family so that these stakeholders may be directly included in the provision of information, allowing them to be better informed, prepare along with the patient, and be involved in recovery planning.
Discussion

Principal Findings

Our findings demonstrate that engaging patients and caregivers in research through the formation of an advisory panel yields a rich source of usable information to guide the development of an mHealth app for the perioperative period of cardiac surgery. Advisory panel members generated 7 key design-specific messages centered around access, tracking, synchronization, and reminders, as well as 16 key content-specific messages centered around medical terms, professional roles, cardiac surgery procedures and recovery, educational videos, travel, nutrition, medications, resources, and physical activity. These findings are novel because despite the increased recognition of the importance of involving patients in research, patient engagement remains underutilized in many health research areas, including mHealth design [9] and cardiac surgery. Further, while patient input is more regularly sought in the commercial technology arena, it is often obtained through focus groups or pilot testing aimed at gathering proprietary data; it is rare that patients and caregivers are engaged as partners and cocreators of mHealth.

Several characteristics of our patient engagement activities likely contributed to the gathering of useful information. The first is the deliberate intention to create an environment that supported patients’ and caregivers’ integration into research through activities that targeted the guiding principles that underlie patient engagement [11] and as led by a skilled facilitator. Second, a mixture of broad and focused open-ended questions was used to gather spontaneous feedback as well as feedback related to categories of information based on our previous work. Interestingly, during discussions of the broad, open-ended questions, topics raised tended to concern the potential benefits of the mHealth app. For example, some of the topics raised by the panel included the technology’s potential to change how patients and caregivers interact with information to better support patient engagement with their health care plan (eg, through the ability to access information ahead of an appointment to prepare questions or know what to expect, by allowing them to fact-check what they thought they heard during appointments without having to rely on outside sources like internet searches) and the potential for caregivers to become more involved in the patient’s journey. Discussions of more focused questions produced key messages more directly related to the design and content of the mHealth app. Third, advisory panel members were selected based on whether they had undergone cardiac surgery within the past 2 years, thereby ensuring accurate recall of their experience and elaborating on the information they did and did not receive as part of their patient-provider interaction. This would have had a positive impact on their abilities to contribute to conversations. Fourth, the advisory panel met on multiple instances, which allowed advisory panel members to reflect on the study questions and their experiences alone or with caregivers and other individuals who supported them during their patient journeys and then to bring these reflections back to enrich discussions in subsequent meetings. Finally, the advisory panel included both patients and their caregivers, which provided a breadth of experiences, and turned out to be timely, given the patients’ statements on the potential of the mHealth app to allow caregivers to be more involved in the patient’s journey.

With the increase of older adults being offered cardiac surgery, there is an urgent need to provide a high level of patient-centered value and quality in our perioperative management. The use of evidence-based ERPs has resulted in more rapid and optimal recovery than that with traditional perioperative methods (ie, improved survivorship) in patients who have undergone cardiac surgery [26]. Although published guidelines provide an important framework from which to develop clinical pathways [27], implementation remains challenging, and therefore, the protocols are underutilized. It is anticipated that the approach of involving patients and caregivers in the development stage will enable the health care team to focus on patient-caregiver value in the subsequent implementation phase that will ideally translate to a sustainable process. To this end, the findings from this study have provided a deeper understanding of patient and caregiver needs pertaining to information delivery about various aspects of perioperative care and the potential role of mHealth in supporting these recommendations.

Limitations

This study has some limitations that warrant mention. Logistical constraints shaped our patient engagement approach. For example, while we engaged patients and caregivers at specific time points within the study, we did not continually involve them throughout the project as full research coinvestigators. Had there been continual engagement, there would have been other points of input and the nature of advisory panel members’ relations with the study would have been different. That said, it is important to note that advisory panel members were invited to be coauthors on this manuscript, both to further support the establishment of authentic research partnerships and to ensure that the manuscript accurately reflects their voices and ideas. We also plan to engage advisory panel members further in the reevaluation and revision of the mHealth app prior to its adoption as a standard of care tool to be used within the Cardiac Sciences Program at St. Boniface Hospital.

Conclusions

In an era of increasingly utilized mHealth technologies for optimizing health care delivery, we demonstrated that patient engagement may successfully facilitate the development of an mHealth app whose design and content are driven by the lived experiences of patients who have undergone cardiac surgery and their caregivers. The result was a detail-oriented and patient-centered mHealth app that helps to empower and inform patients and their caregivers across the perioperative period of cardiac surgery. Applications of different patient engagement approaches and their effects on mHealth app development, measures of feasibility, and health outcomes warrant further study.
Acknowledgments

We acknowledge the contributions of RCA and ASHS as cosenior authors on this manuscript, as well as of AMC and ASHS for bringing expertise in patient engagement approaches to the larger research project and this manuscript. We would like to thank the members of the Healthy Heart Patient Researchers group for their partnership in this research endeavor, which include Ms Carole Dobson, CH, Mr Wayne Dobson, and 7 other members. Further information about this study and our other work in the area of patient engagement in research can be found at [28]. This work was supported by a Geographical Full-Time (GFT) research grant from the Department of Surgery (University of Manitoba). Funding for the development of the study’s mobile health app included in-kind contributions from BeeWell Health. The authors had full control over the data and the company has not seen a version of this publication prior to submission. AMC’s postdoctoral fellowship was supported by the Department of Family Medicine (University of Manitoba) and a Patient Oriented Research Awards-Transition to Leadership Stream-Phase 1 award (Canadian Institutes of Health Research, grant #170670).

Conflicts of Interest

RCA reports grants from Pfizer Canada Inc as well as other (ie, honoraria) from Mallinckrodt Pharmaceuticals, Abbott Nutrition, and Edwards Lifesciences, outside the submitted work. None of the authors have any formal relationship (financial or nonfinancial) with the industry partner in this study. None of the other authors have any other conflicts of interest.

Multimedia Appendix 1

An overview of our study’s approach to creating an environment conducive to patient engagement.

References


5. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization Global Observatory for eHealth. URL: https://apps.who.int/iris/handle/10665/44607 [accessed 2021-11-08]


**Abbreviations**

ERP: enhanced recovery protocol

mHealth: mobile health

©Anna M Chudyk, Sandra Ragheb, David Kent, Todd A Duhamel, Carole Hyra, Mudra G Dave, Rakesh C Arora, Annette SH Schultz. Originally published in JMIR Perioperative Medicine (http://periop.jmir.org), 30.11.2021. 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.
Mobile Health Apps That Act as Surgical Preparatory Guides: App Store Search and Quality Evaluation

Naga Sindhura Gadde¹, MBBS; Kevin Yi-Lwern Yap¹,², PhD

¹Department of Public Health, School of Psychology and Public Health, La Trobe University, Melbourne, Australia
²Department of Pharmacy, Singapore General Hospital, Singapore, Singapore

Corresponding Author:
Kevin Yi-Lwern Yap, PhD
Department of Pharmacy
Singapore General Hospital
SingHealth Tower, 10 Hospital Boulevard, Lobby A, Level 9
Singapore, 168582
Singapore
Phone: 65 63214366
Email: kevin.yap.y.l@sgh.com.sg

Abstract

Background: Mobile health (mHealth) apps are becoming increasingly common in surgical practices for training, education, and communication. Factors leading to increased delays, morbidity, and mortality in surgery include inadequate preoperative patient preparation due to a failure to identify patients and procedure details, and missing instruments and equipment required for the procedure. Many apps are available for supporting preoperative, intraoperative, and postoperative care. However, there is a lack of studies that assess the quality of apps that act as surgical preparatory guides.

Objective: The aim of this study is to evaluate the quality of apps that act as surgical preparatory guides for operating room personnel through an in-house quality assessment tool.

Methods: The quality assessment tool comprises 35 questions categorized into 5 sections: (1) engagement (customization, interactivity, target audience; 19 points), (2) functionality (performance, ease of use, navigation; 12 points), (3) aesthetics (layout, visual appeal; 6 points), (4) information (quality and quantity of information, visual information, credibility; 29 points), and (5) privacy and security (4 points). An app search was conducted in the Australian Apple and Google Play stores using the following keywords: “surgical apps”, “surgical preferences”, “surgeon preferences”, “operating room”, and “perioperative procedures”. The overall total scores and scores for each section were reported as medians and IQRs, expressed as raw scores and percentages.

Results: A total of 5 unique apps were evaluated on both iOS and Android platforms. The median overall score across all apps was 35/70 (50%; IQR 38.6%-64.3%). ScrubUp (48/70, 69%) and MySurgeon (42/70, 60%) had the highest overall scores, followed by PrefCard (35/70, 50%) and Scrubnote (28/70, 40%). The lowest scoring app was BrainPadd (26/70, 37%). The sections with the highest median scores, in decreasing order, were privacy and security (4/4, 100%; IQR 75%-100%), aesthetics (5/6, 83%; IQR 75%-91.7%), engagement (5/19, 26%; IQR 17%-34.5%), functionality (7/12, 58%; IQR 29.2%-75%), and information (5/29, 17%; IQR 15%-34.5%). Most apps scored well (4/4, 100%) on privacy and security, except for Scrubnote (2/4, 50%). ScrubUp received a perfect score for aesthetics (6/6, 100%). MySurgeon (17/19, 90%) had the highest engagement score, while ScrubUp and MySurgeon had the highest functionality scores (9/12, 75% each). All apps scored below 50% for the information section, with ScrubUp having the highest score of 13/29 (45%).

Conclusions: ScrubUp and MySurgeon had the highest quality scores and can be used as adjuncts to hospital protocols by operating room personnel for their surgical preparation. Developers are encouraged to develop appropriate apps for surgical preparation based on relevant guidelines and standards, as well as the quality evaluation criteria in our tool. Operating room personnel can also use this tool as a guide to select and assess their preferred apps in their practices.


KEYWORDS
mHealth apps; surgical apps; surgery preparation; operating room personnel; quality assessment; quality evaluation; perioperative; operative; mobile health; surgery; post-operative
Introduction

As the global digital health market continues to flourish, technological innovations such as electronic medical record systems, laboratory and clinical information systems, mobile apps, health information technology, wearable devices, telehealth, and telemedicine are becoming more common in the health care setting to improve health service delivery and quality [1]. Mobile health (mHealth) technologies such as mobile phones and patient monitoring devices are becoming increasingly common in medical and surgical practices for training, education, and communication [1,2]. Surgical processes are complex, and a variety of factors may result in surgical delays and cancellations [3]. Major factors leading to increased delays, morbidity, and mortality in surgery include inadequate preoperative patient preparation resulting from a failure to identify patients and procedure details, and missing instruments and equipment required for the procedure [4]. Similarly, perioperative nurses face pressures to balance between maintaining operating room schedules and the surgeon's demands due to time restrictions and the potential to underestimate the time required to operate [5]. Furthermore, newly graduated nurses who are uncertain about the intensified and unfamiliar situations during the initial stages of their clinical practice may face additional pressures in addressing emerging and complex technology, resulting in technological stress and surgical errors [6]; these pressures are in addition to trying to cope with the demands of new clinical placements and advancing their professional careers at the same time [7,8].

Surgical preparation tools such as preassessment clinics and the World Health Organization Surgical Safety Checklist aim to minimize adverse events and errors in the operating room [3,9]. However, in some developing countries, surgical teams are still unable to use the safety checklists effectively [10]. With an increasing global trend of smartphone users to a predicted 3.8 billion users in 2021 [11], it is envisaged that smartphone apps may be useful to surgical teams, newly graduated health care professionals, or health care trainees to improve their awareness of and attitudes toward surgical safety practices since these apps can be accessed anytime and anywhere [10,12]. Some features of such apps that may enhance the usability of these safety checklists include the ability to customize surgical preparatory notes according to user preferences, provide comprehensive step-by-step information on preoperative and postoperative procedures, and provide clear and accurate visual explanations of the surgical procedures, for example, through images or videos. Furthermore, smartphone apps can act as quick reference guides that provide various clinical resources as a part of training and education for health care trainees, students, residents, fellows, and surgeons during their clinical practices to ultimately improve communication, system efficiency, and patient safety [1,13-15]. Studies have shown that many apps are available for supporting preoperative, intraoperative, and postoperative care [16], and that health care professionals working in surgical care and app developers are also producing innovative apps to aid surgical teams in education, training, and practice [15].

Despite the widespread availability of mHealth apps, the literature has focused mostly on the prevalence and evaluation of communication, education, clinical, and diagnostic apps for physicians; health and medication monitoring apps for patients; and healthy living apps for diet, exercise, pregnancy, and heart rate monitoring for laypersons [13]. In the surgical domain, some studies have focused on the use of mobile patient health record apps and apps for diagnosis in the perioperative setting (eg, smartphone-based electrocardiograms, pulse oximetry, and blood glucose monitoring), while others have concentrated on the use of medical reference and perioperative crisis event management apps to improve care quality and safety in patient care, as well as on apps that facilitate patient monitoring and follow-up in the postoperative period [17]. The use of smartphones to promote better communication (eg, text messaging, emails) among the care team has also been studied [17]. However, there is a lack of studies that assess the quality of apps that act as surgical preparatory guides. Surgical preparatory guides help to support the surgical team in preparing the operating room and the patients before each procedure. They can consist of checklists of tools and equipment (eg, dressings, drapes, disposables, surgical instruments), information on preoperative safety procedures (eg, checking consent forms and diagnostic images, identification of patients and surgical sites), information about surgical preparatory steps (eg, putting up drapes, sterilizing surgical sites, preparing the patient and operating room, correct positioning of the patient), information on postoperative steps after each procedure (eg, recording correct counts of instruments, procedure names, specimens collected, equipment issues), and related reference sources, among others. Furthermore, there are concerns that there is no control on the quality of the apps, nor is there any regulated body that oversees the validity of the content unless the app is considered a medical device [18]. On the other hand, the lack of proper information on the quality of such apps and their content makes it difficult for users to identify the most useful apps, and there is also a risk that users may access misguided or misleading information [12,19]. Therefore, the objective of this study is to develop a quality assessment tool to evaluate apps that act as surgical preparatory guides for operating room personnel (nurses, surgical technicians, circulating nurses and technicians). The aim is to provide a recommendation of apps that may be useful to operating room personnel during their training and initial stages of clinical practice to improve patient safety and health communication within the surgical team.

Methods

Development of the Quality Evaluation Tool

The overall framework of the quality evaluation tool was adapted from the Mobile Application Rating Scale (MARS) [20], with modifications made to some criteria to fit the evaluation of apps for surgical preparation based on relevant articles found from a keyword search in PubMed (“surgical apps” OR “surgical applications” OR “surgical safety” AND “quality tools” OR “quality scale” OR “assessment criteria” OR “evaluation”) [15,16,18,21-24]. In addition, we used relevant articles that reported quality assessment tools or criteria for evaluating general mHealth apps to refine the quality evaluation criteria of our tool [25-43].
The tool was comprised of 35 questions categorized into 5 sections: (1) engagement (customization, interactivity, target audience; 19 points), (2) functionality (performance, ease of use, navigation; 12 points), (3) aesthetics (layout, visual appeal; 6 points), (4) information (quality and quantity of information, visual information, credibility; 29 points), and (5) privacy and security (security, privacy; 4 points) (Multimedia Appendix 1). The maximum possible score was 70 points.

Selection of Apps
We conducted an app search in August 2020 on the Australian Apple (iOS) and Android (Google Play) app stores with the keywords “surgical apps”, “surgical preferences”, “surgeon preferences”, “operating room”, and “perioperative procedures”, which resulted in an identification of 1110 apps (Figure 1). Among these apps, a total of 642 apps were not related to surgery and were excluded from screening. Among the 468 apps that were screened based on app name and description, a total of 458 apps were excluded based on the following exclusion criteria: education or examination-related, language other than English, and requiring a subscription or payment for access. Surgical-related apps that were not related to surgical preparatory guides were also excluded including communication or coordination tools, apps specific to a hospital or disease, apps targeting non–operating room personnel or patients, game-based apps, journal-related apps, and apps advocating for the purchase of surgical tools. There were 5 unique surgical apps that were evaluated on both iOS and Android platforms. The iOS and Android versions of the apps were evaluated on an iPhone 7 (Apple Inc) and Oppo A37F and X9079 (Guangdong Oppo Mobile Telecommunications Corp., Ltd) phones, respectively.

Evaluation and Data Analysis
The apps were evaluated independently by 5 individual raters. Any discrepancies in scores were resolved by a discussion among all the raters, and the final scores were used for analysis. The iOS and Android versions for each unique app were not treated differently; the results were not reported in relation to the different platforms, but per unique app instead. We computed descriptive statistics for the overall total scores and the scores for each section of the apps using SPSS, version 26 (IBM Corporation) and reported these scores as medians and IQRs, expressed as the raw scores and percentages.

Results
The median overall score across all the apps was 35/70 (50%; IQR 38.6%-64.3%). The apps that had the highest overall scores were ScrubUp (48/70, 69%) and MySurgeon (42/70, 60%), while the lowest scoring app was BrainPadd (26/70, 37%). Among all the apps, only 3 (ScrubUp, MySurgeon, and PrefCard) had an evaluation score of 50% or above (Table 1). The sections with the highest scores, in decreasing order, were privacy and security (4/4, 100%), aesthetics (5/6, 83%), engagement (15/19, 79%), functionality (7/12, 58%), and information (5/29, 17%).

In general, 4 apps took into consideration privacy and security aspects of their features, achieving the maximum score of 4/4 (100%). Only Scrubnote scored 2/4 (50%) as it did not explicitly state any privacy policy. All apps had a security login function for user authentication purposes.
### Table 1. Evaluation scores of the apps.

<table>
<thead>
<tr>
<th>Evaluation scores for the surgical preparation apps, ranked from first (left) to last (right), points (%)</th>
<th>Scores across all the apps</th>
<th>Median (%)</th>
<th>IQR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Engagement (out of 19)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScrubUp&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16 (84)</td>
<td>17 (90)</td>
<td>15 (79)</td>
</tr>
<tr>
<td>MySurgeon&lt;sup&gt;b&lt;/sup&gt;</td>
<td>12 (86)</td>
<td>13 (93)</td>
<td>11 (79)</td>
</tr>
<tr>
<td>PrefCard&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3 (75)</td>
<td>3 (75)</td>
<td>3 (75)</td>
</tr>
<tr>
<td>Scrubnote&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1 (100)</td>
<td>1 (100)</td>
<td>1 (100)</td>
</tr>
<tr>
<td>BrainPadd&lt;sup&gt;e&lt;/sup&gt;</td>
<td>9 (75)</td>
<td>9 (75)</td>
<td>7 (58)</td>
</tr>
<tr>
<td><strong>Customization (out of 14)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScrubUp&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7 (50)</td>
<td>2 (100)</td>
<td>2 (100)</td>
</tr>
<tr>
<td>MySurgeon&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6 (75)</td>
<td>6 (75)</td>
<td>3 (38)</td>
</tr>
<tr>
<td>PrefCard&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>1 (50)</td>
<td>2 (100)</td>
</tr>
<tr>
<td>Scrubnote&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4 (100)</td>
<td>3 (75)</td>
<td>3 (75)</td>
</tr>
<tr>
<td>BrainPadd&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>2 (100)</td>
</tr>
<tr>
<td><strong>Aesthetics (out of 6)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScrubUp&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6 (100)</td>
<td>5 (83)</td>
<td>5 (83)</td>
</tr>
<tr>
<td>MySurgeon&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4 (100)</td>
<td>3 (75)</td>
<td>3 (75)</td>
</tr>
<tr>
<td>PrefCard&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>2 (100)</td>
</tr>
<tr>
<td>Scrubnote&lt;sup&gt;d&lt;/sup&gt;</td>
<td>13 (45)</td>
<td>7 (24)</td>
<td>4 (14)</td>
</tr>
<tr>
<td>BrainPadd&lt;sup&gt;e&lt;/sup&gt;</td>
<td>4 (33)</td>
<td>2 (17)</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Information (out of 29)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScrubUp&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4 (44)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>MySurgeon&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5 (63)</td>
<td>5 (63)</td>
<td>3 (38)</td>
</tr>
<tr>
<td>PrefCard&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>4 (100)</td>
<td>4 (100)</td>
</tr>
<tr>
<td>Scrubnote&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>2 (100)</td>
</tr>
<tr>
<td>BrainPadd&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>2 (100)</td>
</tr>
<tr>
<td><strong>Total score (out of 70)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScrubUp&lt;sup&gt;a&lt;/sup&gt;</td>
<td>48 (69)</td>
<td>42 (60)</td>
<td>35 (50)</td>
</tr>
</tbody>
</table>

Overall, all apps scored well in the aesthetics section (median 5/6, 83%; IQR 75%-91.7%). Only 1 app (ScrubUp) received a perfect score (6/6, 100%), while another app (Scrubnote) scored the lowest (4/6, 67%). In general, the majority of the apps scored well for visual appeal, with consistent colors and fonts, and a clear organization of the content on the screen interface. However, there were some minor difficulties in locating and selecting some icons on most apps, except ScrubUp; thus, these apps only scored 3/4 (75%) in terms of layout.

The median engagement score for the apps was 15/19 (79%; IQR 57.9%-86.8%). MySurgeon (17/19, 90%) had the highest engagement score, followed by ScrubUp (16/19, 84%). BrainPadd scored the lowest in engagement (8/19, 42%). All the apps scored 3/4 (75%) for interactivity as they only allowed one method of feedback about the app. All of the content in the apps was appropriate for their target audiences (1/1, 100%).

In terms of customization (median score 11/14, 79%), all apps allowed users to store personalized notes according to their own preferences and the preferences of their surgical team members. BrainPadd was the only app that did not allow users to edit any preloaded information about surgical preparatory procedures or tools, nor add any additional details or images to the preloaded information, thus scoring the lowest (4/14, 29%). There were only 2 apps that allowed limited customization of notifications (MySurgeon, PrefCard) and syncing of scheduled reminders or alerts (MySurgeon, ScrubUp).

Functionality was the second lowest scoring section, with a median score of 7/12 (58%; IQR 29.2%-75%). While more than...
half of the apps (ScrubUp, PrefCard, BrainPadd) scored well due to their consistency in navigation (2/2, 100%), only 2 apps (MySurgeon, PrefCard) did not have any technical issues (2/2, 100% each). The top scoring apps in this section (ScrubUp and MySurgeon, 9/12, 75% each) also scored the highest for ease of use (6/8, 75% each). These apps enabled users to access saved information without the need for internet access and also provided useful help sections for navigating the apps. By contrast, Scrubnote (3/12, 25%) and BrainPadd (4/12, 33%) scored the lowest for this section.

The information section was the lowest scoring section, with a median score of 5/29 (17%; IQR 15.5%-34.5%). All the apps scored below 50% in this section, with the highest scoring app (ScrubUp) having a score of only 13/29 (45%). PrefCard scored the lowest in this section (4/29, 14%). None of the apps had a preoperative surgical safety checklist or a checklist of postoperative steps to be completed after the surgical procedure. ScrubUp was the only app that had preloaded images of the surgical instruments/tools displayed to users, resulting in a score of 4/9 (44%) in terms of visual information. With regard to credibility, ScrubUp and MySurgeon had the highest scores (5/8, 63%), even though none of the apps provided any information on funding.

**Discussion**

**Principal Findings**

This study analyzed 5 apps that act as surgical preparatory guides for operating room personnel who might need them in their practices, such as those in training or new to the operating room. Among them, 4 apps (ScrubUp, MySurgeon, Scrubnote, PrefCard) could be used by multiple surgical specialties, while BrainPadd was specific to plastic surgery. As surgery procedures are becoming more advanced and complex, the operating room environment needs to be coordinated efficiently through effective communication among the surgical team members [44]. Therefore, the evaluation criteria in the customization section of our developed tool were intended to support the communication of surgical team members in the operating room. Our tool also evaluated the engagement of the apps in terms of customization and interactivity, as well as the user-friendliness of the apps. In general, the engagement scores of all the apps were relatively high. All the apps allowed users to store personalized notes. Users were also able to provide feedback about the apps through at least one form of contact to the company or developer (eg, contact number, email, feedback form). In terms of user-friendliness, the top scoring apps, ScrubUp and MySurgeon, provided useful help sections for users to navigate the apps. Therefore, these apps could potentially be used by the surgical team with little extra training or resources. On the other hand, developers of the other evaluated apps could improve the functionality features, such as providing more detailed instructions or user guides on how to use the apps and providing more obvious navigation links between screens. Furthermore, the importance of checklists was highlighted in several studies as an effective communication tool that could impact operating room efficiency and reduce delays and errors in surgical settings [45-47]. The Surgical Safety Checklist by the World Health Organization was developed to address this challenge of minimizing common and avoidable risks in the operating room before, during, and after the surgery process [9]. Thus, the need for a surgical safety checklist was also evaluated as part of the information section when evaluating the surgical preparatory apps in this study.

Providing evidence-based information is one of the important criteria for medical apps, and this can be said for surgical preparatory apps as well. Studies have shown that health care professionals are more inclined to use apps that can provide current and up-to-date information at the point of care in clinical practice [13]. Similarly, other studies have also reported that users value apps that can provide them with immediate access to information [48]. Our evaluation tool attempted to address these factors in the information section by assessing the up-to-dateness of the app content, as well as the presence of preloaded information and evidence-based references. Unfortunately, none of the apps evaluated in this study contained preoperative surgical safety or postoperative procedural checklists, nor did they provide references for their information. Even though ScrubUp scored the highest in the information section, it only contained checklists of the surgical instruments and tools needed and preloaded images. Interestingly, none of the apps had included any preloaded or linked videos to explain the surgical preparatory procedure. Video-based learning is a useful and effective way of learning about surgical preparation, especially among residents [49]. Although the quality of surgical videos on video sharing sites such as YouTube can be improved, this form of learning presents an opportunity and can be considered for inclusion by app developers, if the videos are accurate, reliable, and evidence-based [50].

In our study, PrefCard was ranked third based on its overall evaluation score. We observed that certain features of the app could only be accessed by users who were on the app’s list of affiliated hospitals. As the raters in our study were not from the list of affiliated hospitals, app features such as customization, functionality, and quantity and quality of the content were evaluated based on the images and descriptions provided in the app stores and the developer website. Other criteria related to functionality, such as performance of the app (eg, technical bugs or crashes), having an autocomplete feature, and accessing saved information in offline mode might have been scored differently from another user who had full access to all the features of the app. Furthermore, during the evaluation of BrainPadd, there were some technical issues with regard to downloading and accessing the app at the later stages; hence, the raters evaluated some criteria based on the description on its website. As with PrefCard, users who have full access to BrainPadd might also score the app differently.

**Limitations**

A limitation of this study was that it only evaluated apps available in the English language and from the Australian app stores. Therefore, our results would need to be extrapolated with caution when applied to apps in other countries. In addition, this study only evaluated the app features that were free. There were some features in the apps that were available as in-app purchases, and these were not included in our evaluation. The
evaluation criteria in our tool were developed specifically to assess apps that were meant to be surgical preparatory guides, and not all surgical apps as the variety of surgical apps was too broad. Thus, evaluators who want to use this tool to conduct their own evaluations of surgical apps would have to modify or adapt the criteria to fit their scope of evaluation. Moreover, usability and user acceptance studies or trials are beyond the scope of this study due to time limitations. Future studies should include evaluating the receptivity and acceptance of these apps among potential users, as well as involving appropriate health care professionals as evaluators, such as those in the surgical team. Lastly, this study did not take into account any updates to the apps after their evaluation in August 2020. Any updated versions of the apps might lead to different scores for the individual sections and the overall quality scores. Users should consider any new or updated features of the apps when interpreting our results.

**Conclusion**

This study, we developed a tool for evaluating apps that act as surgical preparatory guides. Based on our evaluation, ScrubUp and MySurgeon are among the apps with better scoring features and can be used as adjuncts to existing hospital protocols for surgery preparation. In addition, the evaluation criteria in our tool can provide a form of guidance for operating room personnel, surgical professionals, and trainees to evaluate their preferred apps in the future. Similarly, app developers are encouraged to develop apps that follow relevant guidelines and standards, as well as the quality criteria in this tool, so that better quality apps that are reliable and incorporate evidence-based content can be used for surgical practices. Where appropriate, we also encourage app developers to submit their apps to relevant regulatory agencies for further evaluation and feedback.

**Acknowledgments**

The authors would like to thank Dr Michael Liem, Mr Upendra Soppa, Mrs Jyothika Chunduru, and Mrs Vinila Kolla for their help in rating the apps, and Mr Adamm Ferrier, Dr Ali Lakhani, and Mrs Jane Sheats for their advice on analysis and how this study relates to practice.

**Authors’ Contributions**

KYLY conceived and designed the study. NSG conducted the evaluation of the apps and analyzed the results. NSG and KYLY wrote the manuscript. All authors agreed to the publication of the manuscript.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1

Quality evaluation tool developed for surgical preparation apps.

[DOCX File, 28 KB - periop_v4i2e27037_app1.docx ]

**References**


Abbreviations

MARS: Mobile App Rating Scale
mHealth: mobile health

Edited by R Lee; submitted 08.01.21; peer-reviewed by YT Cheung, M Ghazisaeedi, H Karim; comments to author 22.05.21; revised version received 30.05.21; accepted 25.10.21; published 30.11.21.

Please cite as:
Gadde NS, Yap KYL
Mobile Health Apps That Act as Surgical Preparatory Guides: App Store Search and Quality Evaluation
JMIR Perioper Med 2021;4(2):e27037
URL: https://periop.jmir.org/2021/2/e27037
doi: 10.2196/jmir.4927
PMID: 34851296
The Psychological Experience of Frontline Perioperative Health Care Staff in Responding to COVID-19: Qualitative Study

Toni Withiel¹, DPsych; Elizabeth Barson¹, DPsych; Irene Ng²,³, MBBS, M ClinREs, GradCert (CritCareEcho); Reny Segal²,³, MBChB; Daryl Lindsay Goulding Williams²,³, MBBS, MBA, CPI; Roni Benjamin Krieser², MBBS(Hons), M ClinUS, GradCert (CritCareEcho); Keat Lee²,³, MBBS, PG Dip ClinUS; Paul Mario Mezzavia², MBBS, M ClinUS; Teresa Sindoni², BNurs, GradCert (ClinNurs); Yinwei Chen², BNurs, GradCert (ClinNurs); Caroline Anne Fisher¹, DPsych

¹Department of Allied Health, Royal Melbourne Hospital, Parkville, Australia
²Department of Anaesthesia and Pain Management, Royal Melbourne Hospital, Parkville, Australia
³Department of Medicine, University of Melbourne, Parkville, Australia
⁴The Melbourne Clinic, Richmond, Australia

Corresponding Author:
Toni Withiel, DPsych
Department of Allied Health
Royal Melbourne Hospital
300 Grattan St
Parkville, 3010
Australia
Phone: 61 411729045
Email: toni.withiel@mh.org.au

Abstract

Background: The rapid spread of the novel coronavirus (COVID-19) has presented immeasurable challenges to health care workers who remain at the frontline of the pandemic. A rapidly evolving body of literature has quantitatively demonstrated significant psychological impacts of the pandemic on health care workers. However, little is known about the lived experience of the pandemic for frontline medical staff.

Objective: This study aimed to explore the qualitative experience of perioperative staff from a large trauma hospital in Melbourne, Australia.

Methods: Inductive thematic analysis using a critical realist approach was used to analyze data from 9 semistructured interviews.

Results: Four key themes were identified. Hospital preparedness related to the perceived readiness of the hospital to respond to the pandemic and encompassed key subthemes around communication of policy changes, team leadership, and resource availability. Perceptions of readiness contributed to the perceived psychological impacts of the pandemic, which were highly varied and ranged from anger to anxiety. A number of coping strategies were identified in response to psychological impacts which incorporated both internal and external coping mechanisms. Finally, adaptation with time reflected change and growth over time, and encompassed all other themes.

Conclusions: While frontline staff and hospitals have rapidly marshalled a response to managing the virus, relatively less consideration was seen regarding staff mental health in our study. Findings highlight the vulnerability of health care workers in response to the pandemic and reinforce the need for a coordinated approach to managing mental health.


KEYWORDS
COVID-19; perioperative; mental health; qualitative; grief; psychology; health care worker; experience; hospital; trauma; thematic analysis; interview
Introduction

The SARS-CoV-2 (herein referred to as COVID-19) pandemic has profoundly changed the fabric of society and presents unparalleled psychosocial and economic challenges at a global scale. Originally identified in the Chinese city of Wuhan in December 2019 [1], the novel pneumonia has spread rapidly around the world. As of December 3, 2020, the World Health Organization [2] reported 64.2 million confirmed cases of COVID-19 with 1.49 million mortalities. On March 11, 2020, the WHO declared COVID-19 as a pandemic with subsequent lockdown measures implemented in Australia. The state of Victoria in Australia has been significantly impacted by COVID-19 with 74.4% of all cases originating in Victoria as of September 14, 2020. As a consequence of the “second wave” of cases, Victorian hospitals were forced to rapidly enact their emergency protocols.

At the Royal Melbourne Hospital, decision making in response to the pandemic occurred in an expedited fashion, often with limited information available. As a consequence of the pandemic, the hospital shifted to a COVID surge planning strategy, developed from late February 2020. Staff communication avenues for the strategy included a hospital Workplace (corporate social media), dedicated COVID-19 information page (established on February 27, 2020), daily dissemination of information in the heads of department hospital huddle, hospital mangers’ briefings (commencing March 23, 2020), and whole-of-hospital Workplace Chat sessions (commencing April 2, 2020). Personal protective equipment (PPE) requirement guidelines were developed and varied according to information available from the Department of Health and Human Services. These changes were disseminated to staff, as available.

At the frontline of the pandemic are health care workers (HCWs) who play a direct role in the diagnosis, treatment, and care of patients diagnosed with COVID-19. The virulence and mortality rate of the virus, along with depletion of PPE and rapidly evolving workplace roles, place HCWs at a distinct risk of experiencing psychological distress [3]. These risks are further exacerbated by the strict lockdown measures imposed in Victoria, with the psychosocial implications of prolonged isolation remaining largely unknown [4]. Existing literature highlights the psychological vulnerability of HCWs who knowingly jeopardize their own health to uphold the health and well-being of others [5]. Indeed, 3573 Victorian HCWs became infected and a significant proportion of these cases were within the hospital setting [6]. Unsurprisingly, an increasing number of quantitative studies have described elevated incidence of depression, anxiety, insomnia, and distress among HCWs exposed to COVID-19 through their work [3,7-9]. Pooled incidence rates from a recent meta-analysis found elevated levels of anxiety, depression, and insomnia, at levels of 23.2%, 22.8%, and 38.9%, respectively, among HCWs exposed to COVID-19 [7]. Simply being a health care worker has been identified as an independent predictor of psychological distress [3,10]. Additional risk factors have been reported to include being female and being a nurse [7]. Further, contributing factors include speculation around modes of transmission, how quickly the virus has spread, and a lack of definitive treatment protocols or vaccinations [11]. Taken together, it is clear that COVID-19 is having profound mental health impacts on HCWs across a broad range of settings.

To date, the vast majority of studies conducted into the mental health impacts of COVID-19 on HCWs has been quantitative in nature. While affording well-documented strengths in internal validity and replicability, the use of quantitative approaches alone may be insufficient to explore the complex nature of psychological distress during this unprecedented time. A recent position paper summarizing the priorities identified by 24 world-leading experts further reinforced the importance of lived experience in characterizing mental health implications [4]. Hence, qualitative approaches represent an important avenue to characterizing the richness and diversity of HCW experiences.

The limited number of qualitative studies that have been conducted to date have revealed a number of themes. Sun and colleagues [12] examined the psychological experiences of 20 nurses caring for patients with COVID-19 in China. Four themes were generated in that study. The theme “significant amounts of negative emotions at an early stage” highlighted the high degree of exhaustion, helplessness, and unfamiliarity in navigating through the early stages of the pandemic. Self-coping styles was also constructed as a theme with participants describing the adoption of a range of coping mechanisms to psychologically adapt to the pandemic, ranging from avoidance to relaxation and humor. Growth under stress was identified among participant accounts, with many identifying a new found appreciation and gratitude for family, social supports, and health. The final theme reflected the juxtaposition of having both positive and negative emotions occur simultaneously. Similar themes were identified in a qualitative study of 30 frontline health care nurses in Wuhan, China [13], where both positive and negative psychological consequences were described.

Using thematic analysis, Munawar and Choudhry [14] explored challenges and coping mechanisms among HCWs in Pakistan. Participants reported limited media exposure and limiting disclosing of work responsibilities to close others as key coping tools. However, a number of culture-specific coping methods were noted, including religious coping and faith-based practices. Unique challenges around denial by religious scholars were also described, with participants reporting frustration at public noncompliance. Taken together, findings highlight the religious and culture-specific experiences faced by HCWs during the COVID-19 pandemic.

To date, no studies have examined the experience of frontline HCWs in psychologically responding to, and dealing with, the COVID-19 pandemic in an Australian context. This study aims to qualitatively explore the psychological impact of the COVID-19 pandemic among frontline HCWs in the perioperative department of a large, tertiary trauma hospital in Melbourne, Australia. Perioperative teams include HCWs that care for patients before, during, and after a surgery or interventional procedures. Included in this HCW cohort are nurses, theater managers, surgeons, anesthetists, surgical andesthetic trainees, theater technicians, and radiographers. In this area of health care, there is a high throughput of patient turnover,
frequent aerosolizing procedures, and close contact with patients, some of them with unknown COVID-19 status and with aerosolizing-generating behaviors.

**Methods**

**Participants and Setting**

Reported in this paper are the data from the qualitative component of a mixed-methods research study evaluating the psychological impact of the COVID-19 pandemic on perioperative staff. A longitudinal quantitative data collection is ongoing via a 4-weekly survey, and the results of this will be reported elsewhere, after study completion. Participants were frontline HCWs employed in the perioperative department of the Royal Melbourne Hospital. All perioperative staff were eligible to participate. Participants were invited by email to register their interest for the study online via a link on their email at the time points of months 4, 5, and 6 of the 7-month-long quantitative study. All participants who registered “interested” were contacted via their preferred method (email or telephone) and provided with the participant information and consent form about the study. Consenting staff were then scheduled for interviews.

Recruitment for the qualitative arm of this study took place in late August 2020, coinciding with the “tail end” of the second peak of COVID-19 cases in Victoria. At the peak of the second wave (July 30, 2020), cases in the state peaked at 723, with national totals also peaking on the same day at 745 [15]. This was considered an opportune moment to prospectively source participant lived experience, as there was time to reflect on the height of this second peak while still remaining current in participant memory. All study interviews took place between September 1, 2020, and October 6, 2020. The hospital in which the study was conducted was one of the most impacted by COVID-19 in the state with comparatively high numbers of both patients treated and staff infections. The state reached 0 infections on October 31, 2020, and this was maintained for 19 days in a row, marking the end of the second wave.

**Procedure**

One-on-one telephone and video conference interviews were conducted with eligible participants after written consent was obtained. Semi-structured interviews contained 10 items which were a combination of open- and closed-style questions. Questions were adapted from existing literature [12,16,17]. Depending on participant responses, standard prompts or follow-on questions were explored with participants (see Multimedia Appendix 1). Questions were presented in a predefined order, with responses audio recorded to facilitate transcription. Interviews were conducted by CF and EB, both of whom are female, senior clinician psychologists with experience in the conduct of qualitative research. Neither interviewers had worked as direct team members with any respondents, nor had regular clinical involvement with the perioperative department. Interviews were conducted with participants at an agreed upon time and location, with 3/9 (33%) opting to conduct the interview on working-from-home days, and the remainder while on-site at the Royal Melbourne Hospital. All participants consented to interviews being audio recorded. This study was approved by the Melbourne Health ethics committee HREC/63609/MH-2020. Participants were informed of their right to withdraw and advised that all data would be collected and disseminated in a confidential manner.

**Statistical Analysis**

Transcription was completed by an independent researcher (TW) and analyzed using NVivo software (QSR International). To ensure consistency in transcription, an orthographic notation system outlined by Braun and Clarke [18] was adhered to. As such, transcription encompassed “verbatim” responses in addition to nonverbal expression such as laughter and pauses. Cross-checking of 3/9 (33%) of interviews was undertaken by CF to ensure consistency and quality in transcription. The interlistener consistency for the transcriptions was high (4363/4443 words; 98.19% consistency).

Inductive thematic analysis was undertaken to analyze data due to its flexibility to manage small data sets and established guidelines for use [19]. An epistemological approach of critical realism was adopted when exploring and classifying themes, an approach which assumes that participant language reflects lived reality while acknowledging the influence of society and culture on that reality [18]. Codes were identified from transcribed materials which allowed for the establishment of a coding structure. This coding structure was used to cross-code 3 transcripts to ensure reliability [20]. Themes were then collaboratively developed by the research team to ensure findings represented lived experiences [21]. The consolidated criteria for reporting qualitative research guidelines [22] were used to guide reporting in this study.

**Results**

**Participants**

A total of 15 staff registered their interest to participate in the study, with 9 agreeing to participate when contacted directly by the interviewers in the research team. The distribution of staff registering interest was 6, 7, 2, and 0 across the 4 monthly registration opportunities. Thus, the decision was made to cease recruitment after month 4, as it was anticipated all staff wanting to participate would have registered interest by this time. This was also mirrored by the diminishing registration rates as the months progressed.

Participants were 9 staff members employed in the perioperative department of the Royal Melbourne Hospital. Nursing staff (n=4) encompassed scrub nurse, scout nurse, clinical nurse specialist, nurse manager, with staff working across roles in some instances. Medical staff (n=5) included anesthetists, neurosurgeons, and orthopedic surgeons.

The sample included a high proportion of senior, experienced staff, with an average of 19.5 years of clinical experience. There was a fairly even split between gender identity (5/9, 56%) female), with no participants identifying as nonbinary, transgender, or having a different gender identity.
Qualitative Findings

Data saturation was calculated in line with Guest and colleagues [23] with a new information threshold set at ≤5%. Implementing a run length of 2 and base size of 4, data saturation was achieved at interview 6+3. Thematic analysis revealed 4 overarching themes: hospital preparedness to respond, psychological impacts, coping strategies, and adaptation with time. A number of subthemes were also identified and are presented below. The themes and subthemes and their relationships with one another are presented in Figure 1 and are explored with supplementary quotes to assist in interpretation. Quotes have had repetitive words, vocalized filler words (umm, err, ah, etc.), and silent pauses removed for fluency of reading. Parentheses at the end of the quotes provide participant position in the hospital.

**Figure 1.** Visual representation of themes and subthemes.

As seen in Figure 1, several external factors were conceptualized to impact the experience of the pandemic at a personal and team level. On the former, coping strategies were constructed as a major theme in the experience of the pandemic, with subsequent implications for the psychological impacts of the pandemic. Adaptation with time occurred in parallel to external and personal experiences, with change and growth evident in both contexts.

**Hospital Preparedness**

The first emerging theme related to the preparedness of the hospital for the COVID-19 pandemic. This theme related to perception of hospital readiness to face the pandemic and encompassed a number of subthemes illustrated in Table 1. We conceptualized this theme in a top-down fashion, with early changes in hospital policy and communication of these changes having associated implications for patient care. The first subtheme of hospital policy was conceptualized by 3 primary factors: readiness to respond, COVID-centric response, rapid policy change, importance of leadership, and response to mental health. There were differing viewpoints regarding the hospital response preparation, with many finding the rapid changes difficult to manage or navigate. Difficulties with unclear lines of leadership, and a perceived lack in the support for mental health well-being provided were also described.

Under the subtheme of communication, the method of communication and the information sources were the major factors. The methods of communication were reported to be unclear and inconsistent, while departmental meetings and peer consultation were deemed more useful than electronic online communication methods. Altered patient care was the third major subtheme, incorporating the factors of patient implications, changing work roles, and resource availability. Participants were concerned about the impact on care for patients with non-COVID–related health needs. They reported a number of changes to their way of working and projects, concerns about moving areas, and having no work. Comments around resource availability were varied, with several reporting no concerns and others raising issue with the supply and suitability of PPE.
<table>
<thead>
<tr>
<th>Subtheme and factors</th>
<th>Description</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hospital policy</strong></td>
<td></td>
<td><strong>Readiness to respond</strong> Participants’ perceptions of overall hospital preparedness. While some participants felt confident in the hospital’s capacity to respond, others cited concerns.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...it personally made me feel confident that the hospital was well prepared because we got on the front foot early and were preparing for the worst and hoping that it wouldn’t happen [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I always felt like the hospital was always behind the eight ball when it came to particular things [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...in the first wave I found that they were a couple of weeks behind a lot of other hospitals [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>COVID-centric response</strong> Participants reported concern at the COVID-centric nature of hospital policy, with other areas of clinical practice falling to the wayside.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...it was very ICU\textsuperscript{a} centric, very ED\textsuperscript{b} centric they made assumptions that theatre would just not be doing anything....they were very focused on one or two areas instead of the whole organization [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...it didn’t really take into account the need for us to continue to look after non COVID patients [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Rapid policy change</strong> Rapid changes in policy in the first wave were described by many participants which resulted in workplace stress and confusion.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• in the early stages because of the changes happening so rapidly I think it created a lot of a confusion and...there are a lot of people...it was quite stressful [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...processes had to escalate quite quickly and changes and decision making at times[medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Importance of leadership</strong> In response to rapidly changing procedures and policies, participants highlighted the importance of leadership in managing the dynamic workplace and resultant adverse implications of fractured leadership.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...sometimes there were potentially too many cooks in the kitchen with some of the decisions and...who was sort of taking the lead and that I guess that ruffles feathers and it leaves a lot of people confused [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• the leadership across the state and the hospitals have not been clear about what they require from people because most people in health will do whatever you ask as long as they are clear of what you’ve asking and... no one really making a decision because at the end of the day even if the decision is not quite right it’s better to have one than not have one at all [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Response to mental health</strong> While participants reported changes in hospital policy to manage patient flow and procedures, some commented on weaknesses in the hospital’s response to mental health.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ...other organizations did a lot more to bring wellness on site than the Royal Melbourne Hospital did...for instance our EAP\textsuperscript{c} people weren’t on site where as Western Health insisted some commented on weaknesses in the hospital’s response to mental health.</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td></td>
<td><strong>Method of communication</strong> While most participants did feel aware of the hospitals’ response to the pandemic, the methods by which this was communicated was cited as a common source of stress and confusion among teams.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• the only thing that I found in the early days was when we would have other departments come into our area and telling us about changes in PPE\textsuperscript{d} before our leadership had a chance to tell us and I think that was that caused a lot of the early anxiety around things [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• trying to get that clarity about what was happening yeah it was quite challenging at the time [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Information sources</strong> Participants primarily highlighted departmental meetings and informal peer consultation as useful means of information transfer. By contrast, participants reported online platforms (specifically Workplace and the hospital intranet) as less useful.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I was very aware because we were having frequent departmental meetings and updates [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• they used Workplace which I think was a useless tool I think this needed verbal communication not just posting reams of documents on a webpage you know like Facebook people just don’t have the capacity read than and the capacity to even understand what’s changed from day to day [nursing]</td>
</tr>
</tbody>
</table>
### Subtheme and factors

<table>
<thead>
<tr>
<th>Subtheme and factors</th>
<th>Description</th>
<th>Quotes</th>
</tr>
</thead>
</table>
| **Altered patient care** | This subtheme revealed concerns that staff had about changing workplace policy and procedure on patient care during the pandemic. | ...everything is slower you know getting surgery is slower if you’ve got an emergency patient it is slower to get them into the operating theatre and you are worried that they are deteriorating while we are all doing the COVID thing [medical]  
everything is slower you know getting surgery is slower if you’ve got an emergency patient it is slower to get them into the operating theatre and you are worried that they are deteriorating while we are all doing the COVID thing [medical]  
other wards...were looking after patients they’d never had to look after before because the normal wards that look after them had gone hot you know there was so many things happening [medical] |
| **Changing work roles** | While some participants felt their role had not changed substantially during the pandemic, others did communicate changing work roles, with many citing a reduction in work capacity. | we dropped the amount of theatres at the beginning and there was kind of just a lot of people floating around and I think that added to the stress because people were like what if I lose my job what if I get deployed to aged care I haven’t done I haven’t done age care nursing in like ten years and people were just quite unsure [nursing]  
...definitely a lot things that we did went on the backburner [nursing]  
I sat there and twiddled my thumbs and did nothing [medical] |
| **Resource availability** | There was a range of responses in the area of resource availability, with some feeling well-resourced, while others raised concerns about the suitability and availability of PPE | I’ve never been concerned for my ward or the hospital that we will ever have a shortage of PPE [nursing]  
I never at any point at the Royal Melbourne Hospital felt as though I didn’t have what I needed [medical]  
I’ve never ever felt happy with the N95 masks that they’ve had and I remember at the start of this pandemic when we were starting to wear N95 masks and I said to people then these masks do not fit me I do not feel safe in these masks [medical]  
we ran out of N95 masks quite quickly and as an anaesthetic nurse you were like oh I’m in the head end like I like I’m most likely to be exposed to this what if I get it [nursing]  
having the right correct PPE at times it hasn’t always been ideal not necessarily the correct sizes of masks available not the best ideal gowns available [nursing] |

---

**Psychological Impacts of the Pandemic**

This theme related to perceived psychological impacts during the pandemic on self and on colleagues as the 2 major subthemes (Table 2). While many participants did not report needing psychological support themselves, all respondents described seeing psychological impacts in their colleagues. Similarly, many participants reported identifying behavior change in colleagues which was interpreted as being driven by anxiety and stress. The most commonly reported emotion was anger, and other factors identified were the need for psychological support, the spectrum of emotions perceived or experienced, fears of infection, the omnipresent nature of the virus, emotional contagion, and the need for calm.

---

*a*ICU: intensive care unit.  
bED: emergency department.  
cEAP: employee assistance program.  
dPPE: personal protective equipment.
Table 2. Theme: Psychological impact of the pandemic.

<table>
<thead>
<tr>
<th>Subtheme and factors</th>
<th>Description</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for psychological support</td>
<td>This theme related to self-perceptions of the need for psychological support. Most participants reported not requiring nor actively seeking psychological support.</td>
<td>• …it didn’t get to a point where ahh I thought I needed it [medical] • …I just felt like maybe other people needed it more so I just avoided it [nursing]</td>
</tr>
<tr>
<td>Spectrum of emotion</td>
<td>There were a range of emotions experienced by staff which occurred on a spectrum.</td>
<td>• I was just surprised at how upset everyone was by it all [medical] • I haven’t at any point felt particularly stressed [nursing]</td>
</tr>
<tr>
<td>Anger</td>
<td>Interestingly, the most commonly expressed emotional response to the pandemic by participants was anger. Of note, the anger originated from causes external to immediate teams.</td>
<td>• I started out being sort of angry and agitated about things and you know [medical] • I don’t think I was stressed I was angry I still am angry [medical] • I had a couple of weeks where I was just like really angry really really angry all the time and its hard not to bring that home [nursing]</td>
</tr>
<tr>
<td>Fear of infection</td>
<td>A number of participants described a fear of catching COVID-19, mostly in relation to bringing the virus home to family and friends.</td>
<td>• …that was quite draining I think everybody early on just felt shattered at the end of every day because they were mindful of the fact they could be infected at any point we didn’t know which patients were likely to be infected or not infected it was a mystery at that level and you know the kind of weight on you personally that you might get infected might get sick might bring it home to your family [medical] • the main thing that I was worried about was actually bringing the virus home I mean that worried me quite a lot [medical]</td>
</tr>
<tr>
<td>The omnipresent nature of COVID-19</td>
<td>The all-present nature of the virus was cited as a contributing factor to overall emotional state.</td>
<td>• I used to come home and like literally just strip off at the back door and then run for the shower you really needed that bit of a wind down because COVID was following you all the way home [medical] • there was a few months back where I had you know extended family members calling me every day to say ‘are you okay we are really worried about you are you okay’ and sometimes the constant questions even though they were coming from a good place I was trying not to think about it when I went home and trying to sort of switch off [nursing]</td>
</tr>
<tr>
<td>Emotional contagion</td>
<td>Bridging the impact of self and colleagues was the subtheme of emotional contagion. This described the spreading of fear among colleagues which had secondary implications on participants mental health.</td>
<td>• while I was waiting I was just having a conversation with one of the nurses over in theatre and she was kind of talking about the numbers overseas and that in two hundred people working in peri-op one of us was going to die and it was just not something that I needed to hear at 3am and as my mind wasn’t even sort of thinking about that kind of thing I was trying not to [nursing] • there was nervousness in the air... [nursing] • they were basing it on a New York situation which we weren’t going to be like New York or Italy and that was very clear from the outset but the panic that ensued through groups of people made it very difficult for people to think rationally and put together a plan for theatre [nursing]</td>
</tr>
</tbody>
</table>
Subtheme and factors | Description | Quote
--- | --- | ---
Need for calm | In response to emotional contagion, several participants described the need for a calming influence in the workplace, a role which many spontaneously adopted themselves. | • I kind of almost took on a role of being quite calm about it all you know because we can’t all be sort of running around and panicking [nursing]
• I sort of after two weeks had just had enough and told my nursing staff that they were not to participate in the panic with medical staff anymore or the executive and we had to sit down and put a plan into place that was actually going to lead our staff and function through the pandemic [nursing]

Impact on colleagues

Behavioral impacts on colleagues | Many participants identified behavioral changes in colleagues as the primary indicator for distress. | • I have felt that some of my colleagues have gone off the rails a little bit in various ways [medical]
• struggling to communicate with various tasks or even just expressing how unsure they were with the masks...you could tell the nervousness like a bit fidgety like a bit sort of racing around sort of thing [nursing]
• this pandemic that has just knocked people out so now you’ve got low energy low morale [nursing]
• reading between the lines in terms of what they are posting online social media what they are texting and a little bit about their behaviour at work...I suspect that some of them have decompensated a little bit [medical]

Anxiety | Participants described observing stress and anxiety among colleagues | • you could see people were getting stressed the anxiety was definitely building for staff [nursing]
• there was some people that I felt were quite overwhelmed or looked a bit anxious at times [nursing]
• the anaesthetist probably whose anxiety has been probably not managed well and been extremely high I think has reflected of the lack of recognition of the risk in their workforce group [nursing]

Perceived inequality | Perceived inequality between professional groups and wards was cited as a contributing factor to emotion among a team. | • equal recognition is important just because you weren’t in a hot ward it didn’t mean you weren’t working [nursing]
• that’s been really palpable you know through the nurses cause the doctors would say to them ‘ohh why don’t you just work from home just stay at home or just do this were only doing half time here and half time at home’ and its like were not allowed to they don’t see that they have a privilege that others don’t have [nursing]

Coping

The theme of *coping* described the various coping strategies used by participants to manage the pandemic (*Table 3*). These were conceptualized as being either external or internal in nature. More participants reported choosing to use external coping tools (eg, exercise, family, and friends) than internal coping (eg, mindfulness, distraction). Also encompassed under this theme were barriers to using coping strategies that had previously worked, largely stemming from restrictions due to lockdown.

https://periop.jmir.org/2021/2/e27166
### Theme: Coping

<table>
<thead>
<tr>
<th>Subtheme and factors</th>
<th>Description</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal coping</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Distraction          | Participants described using distraction as a management tool. | - we’ve had some renovation things going on at home so that was quite distracting [medical]  
- when I went to do some shifts in ICU\(^2\), I found it was a much more controlled environment and because I was learning new skills and doing something it was a good distraction [nursing] |
| Mindfulness          | Use of mindfulness and meditation-based strategies was described by some participants; many using apps to support implementation. Yet other participants reported that these strategies had not been effective. | - I also utilise a lot of online resources a lot of the apps and things that are out there like your smiling mind app and mindfulness and things like downward down dog so that’s the yoga thing but they’ve also got a lot of relaxation things in there as well so I know a lot of apps [nursing]  
- ...things like meditation just haven’t worked for me because they just don’t on a personal level [medical]  
- I had a few weeks there like every night I would come home and I would just sit down and I would put one of those meditation apps on and just lie there for about an hour and just try and calm down and that was quite useful [medical] |
| Mindset              | Adjustment to personal mindset was noted to be a helpful internal coping strategy. | - I think the ups and downs of your day I found you have to have the mindset of just take it day by day...otherwise you do get bogged down with thinking about how long it has been...that’s why I am all surprised when people say it has been six or seven months because I can’t I haven’t looked back to when this really all started [nursing]  
- like most nurses do I’ve just accepted this is how it is and get on with it [nursing] |
| Role of experience   | The role of clinical experience was perceived as a mitigating factor in the experience of stress, with a number of participants highlighting higher levels of distress among more junior staff. | - I think it does come with experience in our environment and knowing where to go to find the information and being a bit more resourceful yourself I think some more junior casual staff probably felt a little bit more at a loss about what’s happening it’s definitely challenging if you aren’t at work for a few days and you come back and your kind of like what’s new what’s happened what’s changed [nursing]  
- there were sometimes more junior staff who were getting a bit stressed about all the changes [nursing] |
| Gratitude            | A number of participants described feeling grateful for having a job in a time of economic turmoil. | - I’m very grateful that I have a job and I’m grateful that even though I’m coming to the hospital I still get to be able to have face to face with people every day which is a thing that a lot of people don’t have so even though you sort of going in to what you feel is often an unsafe environment the fact that you actually get to still see your colleagues and talk with patients and do all that sort of stuff that’s been good [medical] |
| **External strategies** | Entertainment was regularly cited by participants as an external coping tool. | - when I was home I played music more you know watched a lot of Netflix and stuff like that [medical] |
Work as a social outlet

Work was cited as a social outlet. Connection with colleagues played an important role in coping. Similarly, the absence of continued face-to-face contact was cited as a barrier to work engagement.

- I live alone so I found that coming to work was my social life which was actually really good to have [nursing]
- I think a lot of people who you know they probably find coming to work a bit of a relief from being at home all the time [medical]
- being able to go to work and just chatting to your colleagues that’s a that’s a big thing [medical]
- collegiality professional contact is actually quite important and if you just go to work do the work and go home and don’t talk to anybody else its actually rather dull and not very interesting [medical]

Limiting media intake

Limiting media intake was reported as a strategy for coping

- I literally turned off because...I’m a bit of a news person I like to keep up with the news and stuff like that but I would literally turn it off I wouldn’t watch any news I wouldn’t read any newspapers cause It was just total twenty four seven saturation about this [medical]

Exercise

Exercise was cited by many participants as a helpful external strategy for coping during the pandemic.

- try and exercise that’s usually my best way when I’m stressed [nursing]
- I am a bit of a compulsive exerciser...so you know I’ve got a lot of kind of other things that kind of keep me interested outside of work [medical]
- I’ve managed to get the exercise in that I normally would have done I’m probably healthier physically than I’ve been in a long time [medical]

Family and friends

Family and friends played an important role in coping during the pandemic for many participants.

- my coping strategies are I like to stay in touch with people so that’s you know being in touch with family and friends and checking on them [nursing]
- I’ve actually been in contact with a lot more friends and friends from overseas...old friends from school who we haven’t really caught up for ages [medical]

Team response to stress

Strong team bonds and a self-initiated team response were implemented to help manage emotional distress during the pandemic.

- we were really good at looking after our little family here on the ward like we all know each other very well and we know when someone’s not right [nursing]
- Yeah I think it’s bought us together [nursing]
- We tried to implement just fun things at work like you know like whoever’s got whatever colour on gets a prize for the day or whatever you know sort of fun things they’ve started a cahoot morning quiz thing and just a few things like that [nursing]
- I felt that almost I was the you know trying to get some of these people aside and talk as though I was being the counsellor in a way you know just giving people the opportunity to talk and staff like that [medical]

Barriers

Lockdown

The imposed restrictions on movement and activity were cited as a barrier for engaging in previously used coping strategies by a number of participants.

- I’m struggling a bit not being able to see my family [nursing]
- you know obviously living alone I couldn’t see my family members for a long time [nursing]
- unfortunately a lot of coping strategies and things I did have got shut down because of the pandemic [nursing]

Adaptation With Time

The final theme, *adaptation with time*, reflected change and growth over time, and encompassed all other themes. Within this theme, change was described at both a hospital level and a personal level through psychological adaptation and personal growth (Table 4).
Table 4. Theme: Adaptation with time.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital adaptation with time</td>
<td>Many participants reported a rapid adaptation of the hospital’s response to the pandemic over time. This was reported in the context of the preparedness for the pandemic at the second wave relative to the first wave.</td>
<td>• I think as time went on we got kind of trouble shoted and it became a lot smoother now [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• they were weeks ahead dramatic improvement in the response early response to the second wave compared to the first wave [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• pretty rapidly we marshalled a kind of workforce to deal with the various aspects that needed to be done [medical]</td>
</tr>
<tr>
<td>Psychological adaptation with time</td>
<td>Unlike the linear improvement in the hospital’s response, participants described a cycle of emotional change over time.</td>
<td>• so far it’s kind of been a fairly predictable it’s almost like the stages of grief in a way that’s how I feel about it I think that I am kind of at the flat somewhat resigned point now I haven’t reached a point of kind of acceptance I’m sort of in the depression phase a little bit [medical]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I’ve mostly coped well but there are days where I’ve been emotional and not really understood why and I think it’s probably I’ve just not processed what’s going on [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I think I’ve just become a bit more passive in my work and you know not trying to rock the boat too much and although at times I felt that’s been the wrong thing to do its just psychologically for me it’s just a better thing for me to do at the moment not to rock the boat and get angry and upset and just do what you told to do and try and do your best that’s about it [medical]</td>
</tr>
<tr>
<td>Personal growth</td>
<td>A number of participants identified personal growth during the pandemic, with some also reporting changes in life priorities as a result.</td>
<td>• I also feel proud that I am able to deal with this now [nursing]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I’ve had more contact with people since the pandemic than I did before the pandemic because we were very concentrated on our own lives we used to do a lot of travelling and stuff like that so its actually been a bit of a shift in perspective [medical]</td>
</tr>
</tbody>
</table>

Discussion

Principal Results and Comparison to Prior Work

This study explored the psychological impact of the COVID-19 pandemic among frontline HCWs in the perioperative department of a large, tertiary trauma hospital in Melbourne, Australia. Nine qualitative interviews were conducted with frontline HCWs at the tail end of a second peak of infections in the state of Victoria, Australia. The local environment was a health care service with a high level of exposure to the virus, relative to other services in the state and country. This study is important as it aligns with a recent position paper authored by world-leading experts that reinforced the importance of capturing lived experience when characterizing mental health implications [4]. Qualitative analysis indicated 4 main themes from the data: hospital preparedness to respond, psychological impacts, coping strategies, and adaptation with time.

Participants described several emotional states in response to the pandemic, including anger, anxiety, and gratitude, along with psychological concepts including the impact the team and hospital response had on staff emotional well-being and psychological growth. The predominance of anger in our findings is disparate from existing COVID-19 research which has largely found implications for stress, anxiety, and depression [2,5-7]. Yet, findings are consistent with literature from the SARS crisis which highlights anger as a key manifestation of fear and uncertainty [24,25]. Termed as “state anger,” this emotional expression has been purported to be an indicator of underlying emotional distress [24]. It may be that the timing of data collection impacts upon the manifestation of such underlying anger. Notably, participants in our study were interviewed at the tail end of a second and significant wave of infections, with high numbers of infections and associated quarantine requirements among staff. Given this, it may be that anger is an initial emotional reaction experienced by HCWs who are yet to fully process the emotional state underpinning acute experiences. Divergence may also be explained, at least in part, by the professional demographics of HCWs in this study. Unlike Sun and colleagues [12] and Munawar and Choudhry [14], we interviewed a primarily senior and highly experienced sample. It is possible that the nature of these senior work roles and experience of participants lend themselves more intuitively to concern about decision making and resultant implications for staff, rather than the individual experience of distress per se. In support of this notion, several participants in our sample highlighted the unique vulnerability of more junior staff in responding to the pandemic.

Interestingly, despite all participants observing the need for psychological support in colleagues, most indicated that they themselves did not require formal psychological input. Findings are consistent with extant literature, which highlights the perceived stigma around disclosing mental health concerns among health care professionals [26,27]. Reasons cited for non-disclosure have been reported to include anticipated fears of damage to future career prospects and professional standing [28], underpinned by feelings of shame and professional failure [26]. Importantly, Hassan and colleagues [28] found that the reluctance to disclose mental health concerns persisted despite years of clinical experience. Taken together, findings allude to
a pervasive stigma associated with mental health among HCWs and reinforce the need for a hospital-wide approach to managing the same.

Hospital preparedness to respond was cited as a contributing factor to perceived emotional stress among participants. Inadequate PPE was noted as contributing to work-related stress, including the fear of personal infection and subsequent likelihood of spreading infection to family and friends. Findings are consistent with existing literature, with shortages of PPE hypothesized to underscore the onset of mental health symptoms [29]. Similarly, trust in infection control procedures and PPE were shown to be negatively correlated with emotional exhaustion and anger among frontline HCWs during the SARS crisis [30]. PPE shortages have been found globally, with the COVID-19 pandemic surfacing systemic issues of supply in response to overwhelming demand [31]. Findings of our research highlight the need for strengthened coordination and dissemination of appropriate and well-fitted equipment to ensure frontline staff are adequately protected.

Underscoring perceived hospital preparedness was inconsistent communication of rapid policy changes, particularly evident in the early stages of the pandemic. The enormous flow of information was unprecedented and was described as originating from a variety of sources, including informally through news media and communication with colleagues. Even when occurring on more formal mediums such as through the hospital Workplace platform, participants described feeling overwhelmed by the amount of information. Importantly, failures to ensure consistency of messaging resulted in emotional contagion and spreading of misinformation through staff in our study, which further exacerbated mistrust, stress, and anxiety. Regulation of communication under such demanding circumstances is undoubtedly challenging. However, our findings highlighted the importance of early, efficient, and consistent communication through regulated means, disseminated from one distinctive leadership group, in better containing the narrative and minimizing the likelihood of miscommunication.

Participants in our study demonstrated remarkable resilience. Despite barriers to implementation due to lockdown, participants continued to use a range of internal psychological coping strategies and external tools in response to growing workplace demands. On external coping strategies, many participants reported using exercise, entertainment, and work as a social outlet to manage stress. Internal psychological strategies of distraction, mindfulness, and positive mindset were also commonly implemented. Teams also demonstrated clear responses to stress, including the spontaneous adoption of “fun” work activities and associated enhancement of team solidarity. Our findings are consistent with Sun and colleagues [12] who reported both active and passive psychological adjustments in response to the pandemic.

Many participants acknowledged that both the hospital response and their own psychological state were not static during the pandemic, but evolved overtime. This was relative to both the stage of pandemic, regarding chronological duration and infection rates, and the different psychological stages of processing and managing the situation. A recent meta-synthesis of frontline worker perceptions of working in a pandemic highlighted personal and professional growth as an emergent theme [32].

Limitations

This study is not without limitations. Although generalizability is not the goal of qualitative analysis, the sample was taken from a specific team within a single hospital in Melbourne, Victoria, Australia. As such, it may not be transferrable to HCWs from other clinical areas, disciplines, hospitals, or states. The sample size was also small and may not fully reflect the views of all of the perioperative teams. This may be the case particularly for junior and less experienced HCWs in the team who were underrepresented in the self-selected participants. However, this is one of the first studies to qualitatively examine the psychological impact of COVID-19 on HCWs in Australia. Further insights into the psychological response of this team will also be disseminated on completion of a longitudinal survey currently being undertaken.

Several recommendations to assist HCWs’ psychological functioning for organizations managing pandemic situations can be offered from these data. As noted, clear and consistent communication regarding changes to working practices and procedures is likely to facilitate staff feeling informed and assured in the response of the service. Centralized and visual leadership within the organization, particularly when communicating the service response, is also recommended. Visible and available organizationally promoted staff well-being measures are desired by staff, as well as accommodations to working procedures that are as equitable as possible, across professional disciplines. Acknowledgement of the wide variety of psychological responses that people have to stressful pandemic situations may also be advantageous, and a normalization of these responses, so that individuals do not feel unnecessary concerned about their own response, or pressure to respond in a particular way, may be needed.

Conclusions

The psychological impact of stressful and risky situations for HCWs, particularly during periods of prolonged stress, should not be overlooked. While frontline staff and hospitals have rapidly marshalled a response to managing the virus, relatively less consideration was seen regarding staff mental health in our study. Many participants described an emotional response to the pandemic, though barriers in help seeking remain evident. Hospital preparedness was cited as a contributing factor to emotional response, including availability of well-fitted PPE and concerns with communication of rapidly evolving policies. Encouragingly, participants reported using a varied number of coping mechanisms, with adaptation over time evident in response to local barriers. Findings highlight the vulnerability of HCWs in response to the pandemic and reinforce the need for a coordinated approach to managing mental health.
Acknowledgments
We thank all clinicians that supported the conduct of this study and invested their time and effort into making it possible. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Question schedule.

References
Management of Acute Appendicitis in Children During COVID-19 and Perspectives of Pediatric Surgeons From South Asia: Survey Study

Md Jafrul Hannan1*, MS; Mosammat Kohinoor Parveen2*, MPhil; Md Mozammel Hoque3*, MS; Tanvir Kabir Chowdhury4*, MS; Md Samiul Hasan5*, MS; Alak Nandy6*, DA

1Department of Pediatric Surgery, South Point Hospital, Chittagong, Bangladesh
2Department of Pharmacology, Rangamati Medical College, Rangamati, Bangladesh
3Department of Pediatric Surgery, Chittagram Maa-O-Shishu Hospital Medical College, Chittagong, Bangladesh
4Department of Pediatric Surgery, Chittagong Medical College, Chittagong, Bangladesh
5Department of Pediatric Surgery, Dhaka Shishu Hospital, Dhaka, Bangladesh
6Department of Anesthesiology, Chattogram Maa-O-Shishu Hospital Medical College, Chittagong, Bangladesh

*all authors contributed equally

Corresponding Author:
Md Jafrul Hannan, MS
Department of Pediatric Surgery
South Point Hospital
Apt B3, House 72/A, Road 1, Panchlaish
Chittagong, 4100
Bangladesh
Phone: 880 1819345305
Email: jafrulhannan@gmail.com

Abstract

Background: Nonoperative treatment (NOT) of pediatric appendicitis as opposed to surgery elicits great debate and is potentially influenced by physician preferences. Owing to the effects of the COVID-19 pandemic on health care, the practice of NOT has generally increased by necessity and may, in a post–COVID-19 world, change surgeons’ perceptions of NOT.

Objective: The aim of this study was to determine whether the use of NOT has increased in South Asia and whether these levels of practice would be sustained after the pandemic subsides.

Methods: A survey was conducted among pediatric surgeons regarding their position, institute, and country; the number of appendicitis cases they managed; and their mode of treatment between identical time periods in 2019 and 2020 (April 1 to August 31). The survey also directly posed the question as to whether they would continue with the COVID-19–imposed level of NOT after the effect of the pandemic diminishes.

Results: A total of 134 responses were collected out of 200 (67.0%). A significant increase in the practice of NOT was observed for the entire cohort, although no effect was observed when grouped by country or institute. When grouped by position, senior physicians increased the practice of NOT the most, while junior physicians reported the least change. The data suggest that only professors would be inclined to maintain the COVID-19–level of NOT practice after the pandemic.

Conclusions: Increased practice of NOT during the COVID-19 pandemic was observed in South Asia, particularly by senior surgeons. Only professors appeared inclined to consider maintaining this increased level of practice in the post–COVID-19 world.


KEYWORDS

COVID-19; gastrointestinal; pediatric; global surgery
Introduction

Nonoperative treatment (NOT) of acute uncomplicated appendicitis (AUA) refers to treatments involving prescription of antibiotics only, typically those with aerobic and anaerobic coverage of common bacteria commonly found in the bowel. Examples include treatments with amoxicillin and clavulanic acid, piperacillin sodium and tazobactam sodium, or ciprofloxacin hydrochloride and metronidazole hydrochloride. In the Appendicitis Acuta (APPAC) randomized controlled trial [1], intravenous ertapenem was administered for 3 days, followed by 7 days of oral levofloxacin and metronidazole. A combination of intravenous ceftriaxone and metronidazole was used in another study [2].

In South Asian countries, AUA is diagnosed using clinical features, physical examinations, total blood count, and ultrasonograms during outpatient checkups. Complicated appendicitis is considered if there is evidence of appendiceal perforation, abscess formation, localized peritonitis, or lump formation on clinical examinations supported by ultrasonogram. NOT is considered if diagnosis was clinically suspected as appendicitis plus leukocytosis, with confirmation by ultrasonogram. Patients are asked to report to the outpatient department after 7 days of treatment or to the emergency department in case of worsening conditions. NOT is abandoned when no improvement with antibiotics, clinical deterioration, and development of signs and symptoms of complicated appendicitis are observed.

However, treating pediatric acute appendicitis nonoperatively, rather than by open or laparoscopic appendectomy, still elicits great debate [3-5] spawned by the results of controlled trials and observational studies [6-10], reviews [11-14], and meta-analyses [15-17]. The year 2020 has provided a unique opportunity in which some of this opinion bias, particularly against NOT, can be partially removed. In a recent editorial, the pandemic was considered to provide an opportunity for the ultimate trial for NOT [18]. COVID-19, which is caused by SARS-CoV-2, emerged from Wuhan, China, in December 2019. As of November 2021, the infection has spread globally, with over 250 million confirmed cases resulting in over 5 million deaths worldwide [19]. The response to the pandemic has affected all social and economic aspects of life. However, it also has significantly affected health care facilities [20] in several ways, such as reallocation of hospital facilities to accommodate the influx of COVID-19 patients as well as reducing the number of procedures that would normally be carried out as a result of the repurposing of rooms and maintaining of social distancing protocols to mitigate the spread of the virus. Nearly all countries (90%) have experienced such disruptions to their health services, with low- and middle-income countries reporting the greatest difficulties [20]. This effect has particularly strained surgeries, in general [20,21], and, thus, would have a direct effect on treatment strategies for pediatric AUA. The objective of this study was to determine the effect of the COVID-19 pandemic on the approaches of care for AUA in the South Asian countries of Bangladesh, India, Nepal, and Pakistan, and to determine if any permanent changes in opinions of care will result once the effects of the pandemic have diminished.

Methods

The instrument used in this study was a voluntary survey with no added incentives; the survey was administered via SurveyMonkey (Momentive), a leading provider of survey software and tools, and was sent to 200 pediatric surgeons in Bangladesh, India, Pakistan, and Nepal. The invitation to participate was sent via email, mobile SMS, Facebook Messenger, and WhatsApp to physicians on mailing lists and to physicians in all coauthors’ personal networks. In the invitation, prospective respondents were told the length of time the survey would be available, which data would be stored, the data storage location, and for how long data would be kept. No personal information was collected. The purpose of the study was explained, and the investigators were identified. The open survey consisted of 10 nonrandomized questions pertaining to the respondents’ position, type of institute, and country of practice; the number of appendicitis cases they managed nonoperatively—responses were binned within 10% increments. Respondents were able to change their responses prior to submission and were also alerted to any missed or skipped questions. Data were collected over a 2-week period from September 14 to 28, 2020; each participant was given a unique anonymous identifier. The survey questions were conceived and developed by all coauthors, but no statistical validation was performed. The Checklist for Reporting Results of Internet E-Surveys (CHERRIES) checklist [22] for the survey is included in Multimedia Appendix 1. The final question posed to each pediatric surgeon was whether they would maintain their current COVID-19–imposed level of NOT once the effect of the pandemic on their ability to perform surgery subsides. The study was approved by the Ethics Committee of South Point Hospital, Chittagong, Bangladesh (No. Admn/SPH/190/2020).

The survey response data were analyzed using JMP software (version 15.2; SAS Institute Inc). There was no weighting of responses assigned to certain questions. We use matched-pair tests for each respondent and analyzed the results using a Wilcoxon signed-rank test. The analyses were performed ungrouped and grouped by country, type of institute, and position.

Results

A total of 134 out of 200 physicians responded to the survey, representing a 67.0% response rate. There were no missing data. Respondents’ positions, types of institute, and countries of practice are summarized in Table 1.
Table 1. Descriptive statistics of the survey participants’ responses regarding position, hospital type, and country of practice.

<table>
<thead>
<tr>
<th>Participant information</th>
<th>Participants (N=134), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Position</strong></td>
<td></td>
</tr>
<tr>
<td>Resident</td>
<td>14 (10.4)</td>
</tr>
<tr>
<td>Registrar or assistant registrar</td>
<td>21 (15.7)</td>
</tr>
<tr>
<td>Assistant professor</td>
<td>26 (19.4)</td>
</tr>
<tr>
<td>Associate professor</td>
<td>29 (21.6)</td>
</tr>
<tr>
<td>Professor</td>
<td>20 (14.9)</td>
</tr>
<tr>
<td>Consultant</td>
<td>24 (17.9)</td>
</tr>
<tr>
<td><strong>Type of institute</strong></td>
<td></td>
</tr>
<tr>
<td>Private practice</td>
<td>11 (8.2)</td>
</tr>
<tr>
<td>Private medical college</td>
<td>35 (26.1)</td>
</tr>
<tr>
<td>Medical university</td>
<td>12 (9.0)</td>
</tr>
<tr>
<td>Government medical college</td>
<td>44 (32.8)</td>
</tr>
<tr>
<td>Corporate hospital</td>
<td>9 (6.7)</td>
</tr>
<tr>
<td>Children’s hospital</td>
<td>23 (17.2)</td>
</tr>
<tr>
<td><strong>Country of practice</strong></td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>28 (20.9)</td>
</tr>
<tr>
<td>Nepal</td>
<td>6 (4.5)</td>
</tr>
<tr>
<td>India</td>
<td>27 (20.1)</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>73 (54.5)</td>
</tr>
</tbody>
</table>

The change in usage of NOT by each physician in the entire cohort, as well as grouped by position, institution type, and country, is depicted by the heat maps in Figure 1, A to D, respectively. The values in each box represent the percentage of physicians. Boxes that lie completely above the diagonal represent physicians who increased their usage of NOT during the pandemic. The results from the Wilcoxon signed-rank tests are given in Table 2 for the entire cohort as well as grouped by position, institute, and country. There was a significant increase in the percentage of appendicitis cases treated nonoperatively from April 1 to August 31, 2020, compared to the same time period in 2019 ($P<.001$), as clearly observed in Figure 1, A. Though it appears there may be some differences when the cohort is grouped by type of institute or country (Figure 1, C and D), these were not found to be significant ($P=65$ within pairs and $P=.52$ among pairs for type of institute; $P=.65$ within pairs and $P=.52$ among pairs for country). When grouping by position, we found a weak significance within pairs ($P=.06$), which compares the differences between pairs in a group, but not among pairs ($P=.12$), which compares the averages of the pairs between groups.
Figure 1. Heat maps depicting the change in usage of nonoperative treatment (NOT) by each physician in the entire cohort (A) and grouped by position (B), institution type (C), and country (D). Boxes entirely above the diagonal indicate an increase in usage of NOT. The values in each box represent the percentage of physicians (N=134).

Table 2. Summary of statistics from the matched-pair analysis comparing percentage of appendicitis cases receiving nonoperative treatment in 2019 and 2020 for the entire cohort and the grouped by position, institute, and country.

<table>
<thead>
<tr>
<th></th>
<th>P value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Samples analyzed</strong></td>
<td></td>
</tr>
<tr>
<td>Entire cohort</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Within pairs</strong></td>
<td></td>
</tr>
<tr>
<td>Grouped by position</td>
<td>.06</td>
</tr>
<tr>
<td>Grouped by institute</td>
<td>.65</td>
</tr>
<tr>
<td>Grouped by country</td>
<td>.65</td>
</tr>
<tr>
<td><strong>Among pairs</strong></td>
<td></td>
</tr>
<tr>
<td>Grouped by position</td>
<td>.12</td>
</tr>
<tr>
<td>Grouped by institute</td>
<td>.52</td>
</tr>
<tr>
<td>Grouped by country</td>
<td>.25</td>
</tr>
</tbody>
</table>

<sup>a</sup><sup>P</sup> values were based on Wilcoxon signed-rank tests comparing 2019 and 2020 values.

This potentially interesting effect of the position grouping is shown in more detail in the matched-pair plot in Figure 2. As this is grouped data, the y-axis reports the mean of the differences between the groups, and the x-axis represents the mean of the means. For clarity, the data points have been removed and only the group labels are plotted. The dotted red lines indicate the boundaries of the 95% CI, and the solid line represents the mean for the entire cohort. It should be noted that the upper 95% CI value falls below zero. Not only does this plot clearly show that the percentage of appendicitis cases treated nonoperatively has significantly increased in 2020, but it also indicates that professors, consultants, and associate professors were most likely to expand their use of NOT. This is also evident from Figure 1, B.
Figure 2. Matched-pair plot for positions showing the difference between pairs plotted against the average between pairs. The time periods in 2019 and 2020 are April 1 to August 31. The red dotted lines indicate the boundaries of the 95% CI and the solid line represents the mean for the entire cohort. NOT: nonoperative treatment.

Also interesting are the results from the final question as to whether or not the physicians would consider maintaining the COVID-19–imposed level of NOT for appendicitis after the effects of the pandemic on surgery capability subside. An interesting result from this points to the level and spirit of opinion on the topic of NOT, in general. Though the respondents were given the option of replying “don’t know” or “no comment,” the majority (126/134, 94.0%) responded with a definitive “yes” or “no.” The fractions of those grouped by position, institute, and country are shown in the mosaic plots in Figure 3. As in the matched-pair analysis, no association was found using the Fisher exact test between institute ($P=.99$) or country ($P=.37$) and a planned change in the level of NOT practice after the effects of the pandemic subside. However, a significant association was found for position ($P=.04$). The mosaic plot in Figure 3 shows that residents were the least likely, and professors the most likely, to retain their current level of NOT postpandemic.

The data sets generated and analyzed during this study are available online [23].

Figure 3. Mosaic plots illustrating the distribution of responses to the question of adopting nonoperative treatment (NOT) for appendicitis after the effects of the pandemic on surgery capability subside, between April 1 and August 31, 2020.
Discussion

Principal Findings
The results from this study indicate that, despite being required to be more reliant on the usage of NOT for pediatric appendicitis as necessitated by the effects of the COVID-19 pandemic on health care, it has done little to sway the opinion of most physicians in Bangladesh, Nepal, India, and Pakistan regarding its place in the treatment of AUA. Only those practicing with appointed professorships have appeared to give the idea of making the increased usage of NOT a permanent practice after the effects of the pandemic wane. The results of this study indicated an increased usage of NOT, particularly by professors, associate professors, and consultants. This is in good agreement with a recent study by Irish surgeons, where 76% of the 161 participants disclosed that they had modified their practice to a predominantly NOT approach [24]. Similarly, in a New York hospital that traditionally favors operative management, NOT usage increased to roughly 50% of all cases as a result of the pandemic, with favorable results [25]. In an Indian study, Verma et al [26] documented how their institute increased their percentage of NOT to 69% of patients, a significant increase from 22% recorded during the same time period in 2019. The recent work of Ielpo et al [27] also showed an increased usage of NOT in pediatric appendicitis during the pandemic, as did a study from the United Kingdom [28]. More recently, the usage of NOT and the resulting outcomes of AUA in the setting of COVID-19 was systematically reviewed. The results indicated that NOT was a safe, short-term alternative to surgery and led to acceptably low rates of failure and complication [29]. However, a global study on the effect of COVID-19 on pediatric surgery, in general, showed less of an effect, where only 4 out of 20 institutes were found to begin using NOT for appendicitis, and 14 out of 20 reported that they had not changed their extent of NOT use [30].

Another interesting observation coming from this study relates to the results from Pakistan. During the pandemic, there was a clear mandate to change all procedures in children that showed no clinical or radiological signs of complicated appendicitis to a treatment regimen of antibiotics administered intravenously [31]. Yet, the results of this study (eg, Figure 1, D) do not reflect this policy particularly well unless the majority of cases were overwhelmingly complicated appendicitis cases. This may be due to lockdown-related delays in hospital visits of children with acute appendicitis that spawned an increased rate of complicated appendicitis, as documented in a Spanish study [32].

The results of this study indicated that junior physicians were more likely to retain their prepandemic rate of operative treatment compared to their more senior counterparts. While the reason for this is currently unknown, the survey results indicate that 79% of those who identified themselves as residents came from government medical colleges and universities. It might be plausible that the reason for their reluctance to convert to NOT may be a result of the confounding factor that physicians at these institutions are also under pressure to simultaneously maintain their level of education, particularly since there is, in general, a shortage of surgeons in these regions. This would provide a driving force for them to gain experience and continue with surgical techniques.

In all previous studies on the treatment of AUA amid the COVID-19 pandemic, little, if anything, is discussed about how the physicians would modify their practice of NOT once the crisis is over. This study directly addressed this question and the results indicated that it is mainly professors who would consider adapting to change, and residents would be the least likely. The residents would have been in the early stages of the learning curve of surgical techniques, be it open or laparoscopic, and may have had less experience with NOT regimens and the diagnostics required for their safe recommendation. Based on this, it would be understandable that they would not be keen on moving more toward NOT methods when they are in the midst of mastering their surgical skills.

A limitation of this study, which is independent of the physicians’ personal opinions or preferences, is that the resources and capacities of hospitals will be different postpandemic, a factor that was not measured in the survey. This could also sway the usage of NOT toward those facilities with the scarcest of capacities and resources. Ultimately, all of these factors can and should contribute to treatment decisions [33]. Even when the effects of the pandemic subside, the economic effects will lag considerably, and recovery will be slow. This will, in turn, continue to place economic pressures on institutions, which may necessitate longer-than-anticipated usage of NOT for AUA. Another limitation of this study was that the survey was not validated in any formal manner, as it was designed as more of an information gathering tool. Selection bias may be a factor, though the use of both personal networks from all coauthors, rather than just one, and mailing lists used to reach out to potential candidates helped to alleviate this effect.

Conclusions
The global COVID-19 pandemic has resulted in severe disruption of surgeries worldwide; this has necessitated new approaches to maintain care for those in need, when so many resources are being repurposed to address the massive influx of patients stricken by the virus. In the specific area of pediatric appendicitis, there is ample evidence that more hospitals and institutes are increasing their implementation of NOT for AUA in Bangladesh, India, Pakistan, and Nepal. In some respects, pediatric surgeons could think of this as their participation in an involuntary clinical trial. The results suggest that only professors in these countries would consider maintaining this increased level of practice in the post–COVID-19 world and that the effect of institution type or country was insignificant. It is likely that the demands of continuing education of the younger cohort contributes to their reluctance to practice more NOT postpandemic, just as they are beginning the climb the learning curve for operative treatments.
Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

References


Abbreviations

APPAC: Appendicitis Acuta
AUA: acute uncomplicated appendicitis
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
NOT: nonoperative treatment
Predicting Prolonged Apnea During Nurse-Administered Procedural Sedation: Machine Learning Study

Aaron Conway1,2,3, PhD; Carla R Jungquist4, PhD; Kristina Chang2, MScN; Navpreet Kamboj1, MScN; Joanna Sutherland5, MBBS; Sebastian Mafeld6, MBBS; Matteo Parotto7,8,9, MD, PhD

1Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
2Peter Munk Cardiac Centre, Toronto General Hospital, Toronto, ON, Canada
3School of Nursing, Queensland University of Technology, Brisbane, Australia
4School of Nursing, The University at Buffalo, Buffalo, NY, United States
5Rural Clinical School, University of New South Wales, Coffs Harbour, Australia
6Joint Department of Medical Imaging, Toronto General Hospital, Toronto, ON, Canada
7Department of Anesthesiology and Pain Management, Toronto General Hospital, Toronto, ON, Canada
8Department of Anesthesiology and Pain Medicine, University of Toronto, Toronto, ON, Canada
9Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Aaron Conway, PhD
Lawrence S. Bloomberg Faculty of Nursing
University of Toronto
Suite 130
155 College Street
Toronto, ON, M5T 1P8
Canada
Phone: 1 (416) 340 4654
Email: aaron.conway@utoronto.ca

Abstract

Background: Capnography is commonly used for nurse-administered procedural sedation. Distinguishing between capnography waveform abnormalities that signal the need for clinical intervention for an event and those that do not indicate the need for intervention is essential for the successful implementation of this technology into practice. It is possible that capnography alarm management may be improved by using machine learning to create a “smart alarm” that can alert clinicians to apneic events that are predicted to be prolonged.

Objective: To determine the accuracy of machine learning models for predicting at the 15-second time point if apnea will be prolonged (ie, apnea that persists for >30 seconds).

Methods: A secondary analysis of an observational study was conducted. We selected several candidate models to evaluate, including a random forest model, generalized linear model (logistic regression), least absolute shrinkage and selection operator regression, ridge regression, and the XGBoost model. Out-of-sample accuracy of the models was calculated using 10-fold cross-validation. The net benefit decision analytic measure was used to assist with deciding whether using the models in practice would lead to better outcomes on average than using the current default capnography alarm management strategies. The default strategies are the aggressive approach, in which an alarm is triggered after brief periods of apnea (typically 15 seconds) and the conservative approach, in which alarms are only triggered after 30 seconds.

Results: A total of 384 apneic events longer than 15 seconds were observed in 61 of the 102 patients (59.8%) who participated in the observational study. Nearly half of the apneic events (180/384, 46.9%) were prolonged. The random forest model performed the best in terms of discrimination (area under the receiver operating characteristic curve 0.66) and calibration. The net benefit decision analytic measure was used to assist with deciding whether using the models in practice would lead to better outcomes on average than using the current default capnography alarm management strategies. The default strategies are the aggressive approach, in which an alarm is triggered after brief periods of apnea (typically 15 seconds) and the conservative approach, in which an alarm is triggered for only prolonged periods of apnea (typically >30 seconds).

Conclusions: Decision curve analysis indicated that using a random forest model would lead to a better outcome for capnography alarm management than using an aggressive strategy in which alarms are triggered after 15 seconds of apnea. The model would not be superior to the conservative strategy in which alarms are only triggered after 30 seconds.
procedural sedation and analgesia; conscious sedation; nursing; informatics; patient safety; machine learning; capnography; anesthesia; anaesthesia; medical informatics; sleep apnea; apnea; apnoea; sedation

Introduction

With the recent increase in the use of electronic monitoring devices in the hospital setting, alarm fatigue has become a serious problem that impacts patient safety and nursing care [1]. Alarm fatigue is caused by exposure to excessive and frequent device alarms and leads to desensitization to alarms. Alarm fatigue has been linked to patient deaths resulting from delayed responses to clinical deterioration by clinicians who have become desensitized to alarms [2]. One of the sources of alarms is the capnography device that is used to measure and monitor ventilation in patients. A capnography waveform displays the level of expired carbon dioxide (CO₂) over time to show changes in concentrations throughout the respiratory cycle. Capnography waveform abnormalities assist in detecting and diagnosing specific conditions such as partial airway obstruction and apnea. For this reason, the implementation of capnography into practice for respiratory monitoring is considered a high priority to improve patient safety by leading authorities, including national and international professional organizations for anesthesiology in Canada, the United States, and Europe [3-5]. Capnography is commonly used for nurse-administered procedural sedation [6-8], including in the interventional radiology setting [9-13].

Distinguishing between the capnography waveform abnormalities that signal the need for clinical intervention for an event and those waveform abnormalities that do not indicate the need for intervention is essential to the successful implementation of this technology into practice. For example, alarms triggered after short periods of apnea lead to frequent interruptions and potentially increase the risk of alarm fatigue. Conversely, intervention provided only when an apneic event reaches a longer threshold negates the potential benefits that capnography can have on patient safety through improved ventilation. In practice, two alternative strategies for capnography alarm management are typically used. The aggressive strategy involves alarms triggered after short periods of apnea (typically 15 seconds). The conservative approach involves alarms triggered only when an apneic event is prolonged (typically >30 seconds). Preferences for the aggressive or conservative alarm threshold are influenced by many factors, including the rate of oxygen supplementation. The duration of time between the onset of apnea to hypoxemia increases with higher oxygen flow [14].

Capnography alarm management may be improved by using machine learning to create a “smart alarm” that can alert clinicians to apneic events that are predicted to be prolonged. Such an approach aligns with a call from The Society for Critical Care Medicine Alarm and Alert Fatigue Task Force that machine learning techniques should be used to advance the quality of alerts that clinicians receive and to individualize alert delivery based on clinician response characteristics such as alert frequency and event severity [15].

In the aggressive alarm management strategy, if an alarm is only triggered for apneic events predicted to be prolonged, it would reduce the total alarm burden and potentially reduce the risk of alarm fatigue. The downside of applying a machine learning approach to the aggressive strategy would be that some patients with prolonged apnea may not receive early intervention if the model incorrectly predicts that the apneic event will not last for >30 seconds (ie, false negatives). In the conservative alarm management strategy, if an alarm is triggered at the 15-second timepoint for apneic events predicted to be prolonged, it could reduce the total time of the apneic event because treatment could be initiated earlier. The downside of applying a machine learning approach to the conservative strategy would be the potential increase in the total alarm burden if the ratio of false positives (apnea incorrectly predicted to last for >30 seconds) to true positives (apnea correctly predicted at the 15-second time point to last for >30 seconds) is high. This study aimed to determine the accuracy of machine learning models for predicting at the 15-second timepoint if an apneic event will persist for >30 seconds. This information would help determine whether operationalizing these predictions into practice as alarm triggers would be beneficial.

Methods

We performed a secondary analysis of a prospective observational study. The primary aim of the observational study was to identify common patterns in capnography waveform abnormalities and factors that influence these patterns [16]. All participants provided written informed consent and the study was approved by human research ethics committees (UCH HREC 1614; SVHAC HREC 16/26; QUT 1600000641).

Prediction Goal

The prediction goal was to classify apneic events at the 15-second timepoint as either short (ie, terminated before 30 seconds) or prolonged (persisted for >30 seconds). The prediction algorithm was compared against typical default alarm settings for capnography monitors.

Participants

Participants in the observational study were consecutive adult patients who were scheduled to undergo an elective procedure in the cardiac catheterization laboratory with moderate sedation. Patients with severe cognitive impairment who could not provide informed consent and those unable to understand and speak English (in the absence of an interpreter) were excluded. Data collection was performed at two urban private hospitals in Australia.
Sedation and Monitoring

The sedation regimen used for patients included in this study comprised bolus doses of intravenous midazolam and fentanyl. Sedation was performed by nurses who were trained in advanced life support. Routine clinical monitoring included continuous cardiac rhythm and oxygen saturation monitoring as well as noninvasive blood pressure measurements every 5-10 minutes. The Respironics LoFlo Sidestream etCO₂ sensor was used for capnography monitoring. A CO₂ sampling cannula was inserted into the side port of an oxygen face mask or was integrated as a separate line for nasal cannulas. The capnography waveform was displayed on the main physiological monitoring screen. A default “No breaths detected” alert was triggered for apnea, but no other audible or visual alarms were set for the capnography monitor. No restrictions or specific instructions regarding the detection of capnography waveform abnormalities were provided to clinicians as part of the research protocol because the study used an observational design.

Data Collection

Data were collected from August 2016 to May 2018. Demographic data and clinical characteristics were collected from medical records or directly from participants prior to procedures. Intraprocedural data were collected in real time by the researcher present in the procedure room. Direct observation of the participant was required to record the timing of sedation administration and any interventions by sedation providers.

Predictor Variables

Several raw demographic (age, sex) and clinical (American Society of Anesthesiology physical status classification, diagnosis of sleep apnea, BMI, dose and type of sedative and analgesic administered) variables were used as predictors. Features related to sedation regimen dosing used as predictors in the model were the total sedative dose and number of sedative doses administered, time since first sedation, and time since the previous sedative dose. Other features were extracted from the capnography waveform for use as predictors, such as the previous respiratory state (normal or abnormal breathing), duration of the previous apneic event, time since the previous apneic event, and total number of apneic events. A total of 18 predictor variables were used.

Statistical Analysis

Analyses were performed using R version 4.0.3 [17]. Data as well as details about how to access the code and a reproducible computing environment to verify the results were available [18,19].

Modeling

We selected several candidate models to evaluate, including a random forest model, generalized linear model (logistic regression), least absolute shrinkage and selection operator regression, ridge regression, and the XGBoost model. Out-of-sample accuracy of the models was calculated using 10-fold cross-validation. Many participants in the study contributed multiple apneic events to the dataset used for modeling. To take this dependency into account, we ensured that apneic events from individual participants were not included in both the training and testing partitions of the 10-fold cross-validation process. Preprocessing steps included normalizing numeric predictors and using an interaction term for the duration of the previous respiratory state and the total number of apneic events. The discriminatory ability of the models was compared using the area under the receiver operating characteristic curve (AUROC) as well as by plotting sensitivity, specificity, positive predictive values, and negative predictive values (termed a threshold performance plot). A calibration plot with locally estimated scatterplot smoothing was used to assess calibration [20]. The runway package was used to create the plots [21].

Decision Curve Analysis

We used the net benefit decision analytic measure to assist with deciding whether using the models in practice would lead to better outcomes on average than using the current default capnography alarm management strategies. The default strategies are the aggressive approach, in which an alarm is triggered after brief apneic events (typically 15 seconds), and the conservative approach, in which an alarm is triggered for only prolonged apneic events (typically >30 seconds). Calculation of the net benefit essentially transforms the total number of true positives (apneic event predicted to be prolonged at 15 seconds and correctly persisted for >30 seconds) and false positives (apneic events predicted to be prolonged at 15 seconds but did not persist for >30 seconds) into a standardized scale, weighted by the relative harm of a false-positive result [22]. For example, a net benefit of 0.07 means that the net benefit of using the model would be 7 true positives from every 100 predictions from the model. This net benefit can result from any combination of true positives and false positives [23]. A probability threshold of 0.5 indicates that avoiding a false positive is as important to a clinician as identifying a true positive. Preferences for probability thresholds below 0.5 are weighted such that identifying a true positive is more valuable than avoiding a false positive. Preferences for probability thresholds above 0.5 are weighted such that avoiding a false positive is more valuable than identifying a true positive. For example, for a probability threshold of 0.75, the value of a false positive is 3 true positives (0.75/0.25). In other words, to create a net benefit from using the model at this probability threshold, there must be more than 3 true positives for every false positive prediction made from the model. Conversely, for a probability threshold of 0.25, the value of a false positive is weighted far lower, at only one-third of a true positive (0.25/0.75). This means that a net benefit would be achieved if there were more than 1 true positive for every 3 false positives. Decision curves can be interpreted such that the strategy with the highest net benefit at each probability threshold has the highest clinical value [23].

We created a decision curve to plot net benefits across a range of probability thresholds for the aggressive strategy (alarm triggered at 15 seconds of apnea) and the conservative alarm management strategy (alarm triggered at 30 seconds of apnea). The decision curve takes into account the full range of reasonable clinician preferences for the point at which an alarm should be triggered to signal an apneic event in a patient. We tested thresholds in the range of 0.3-0.5 for the aggressive
strategy. In the practical sense, this means that we decided that all clinicians who usually use the aggressive strategy would not accept a probability of prolonged apnea lower than 0.3 as a useful alarm trigger because there would be little difference between this strategy and simply setting the alarm for all apneic events. We also decided that all clinicians would always consider that an alarm is triggered for the aggressive strategy if the probability of prolonged apnea was higher than 0.5. A range of values was used because these probability thresholds can be interpreted as value preferences that individual clinicians may reasonably choose in clinical practice. For example, a clinician who is more risk-averse may select a more conservative probability threshold (closer to 0.3). Individual participant characteristics will also influence clinicians’ decisions about probability thresholds. A clinician may elect to intervene when the probability of prolonged apnea is 0.3 for an older patient with multiple comorbidities but not for a young patient who may more likely be able to tolerate longer periods of apnea. For the conservative strategy, we chose to plot the range of probability thresholds from 0.7 to 0.8. Higher values were chosen because the number of false positives would be an important consideration for clinicians already using a conservative alarm management approach.

**Results**

A total of 384 apneic events of at least 15 seconds duration from 61 of the 102 patients (59.8%) who participated in this observational study were included in the present analysis. A summary of participant characteristics is presented in Table 1. Nearly half of the apneic events (180/384, 46.9%) were prolonged (ie, >30 seconds).

**Table 1.** Participant characteristics (N=61).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR)</td>
<td>76 (68-80)</td>
</tr>
<tr>
<td>Body mass index, median (IQR)</td>
<td>26.4 (24.6-29.8)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>22 (36)</td>
</tr>
<tr>
<td>Male</td>
<td>39 (64)</td>
</tr>
<tr>
<td><strong>Obstructive sleep apnea, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>46 (75)</td>
</tr>
<tr>
<td>Yes</td>
<td>15 (25)</td>
</tr>
<tr>
<td><strong>American Society of Anesthesiology physical status classification, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>I or II</td>
<td>37 (61)</td>
</tr>
<tr>
<td>III or IV</td>
<td>24 (39)</td>
</tr>
</tbody>
</table>

**Discrimination**

A plot of the AUROC for the models using predictions from the 10-fold cross-validation is presented in Figure 1. The random forest model had the best discriminatory power of the models, with a mean AUROC score of 0.66 (SE 0.03). A threshold performance plot, which summarizes the discriminatory power for the models, including values for sensitivity, specificity, positive predictive value, and negative predictive value across all probability thresholds, is presented in Figure 2.
Figure 1. Area under the receiver operating characteristics curve. LASSO, least absolute shrinkage and selection operator.
Figure 2. Threshold performance plot for all models evaluated. LASSO, least absolute shrinkage and selection operator; PPV, positive predictive value; NPV, negative predictive value.

Calibration

The random forest model had the best calibration. It approximated observed risk at moderate (0.5) to high (0.8) thresholds (Figure 3), although the risk was overestimated at very low thresholds and slightly underestimated between 0.4 and 0.5. Other models severely overestimated risk at low probability thresholds and underestimated risk at high probability thresholds.
Figure 3. Calibration plot for all models evaluated. LASSO, least absolute shrinkage and selection operator.

Decision Curve Analysis
As the random forest model performed the best in terms of discrimination and calibration, we chose this model for evaluation using decision curve analysis. The net benefit associated with the random forest model exceeded that associated with the aggressive strategy across all probability thresholds in the range of 0.3-0.5 (Figure 4). The interpretation is that the best clinical outcome would be achieved for clinicians who are willing to initiate intervention for apnea at the 15-second mark if the probability of the event being prolonged was more than 40% by using the random forest model. The net benefit associated with the random forest model was lower than that associated with the conservative strategy across all probability thresholds in the range of 0.7-0.8 (Figure 4). Figure 4A is the comparison of the model with the aggressive strategy and Figure 4B is the comparison of the model with the conservative strategy.
In this study, we found that the random forest model had the best discriminative ability and calibration for predicting if an apneic event would be prolonged during nurse-administered procedural sedation. However, it should be noted that the accuracy of this random forest model was still quite low (AUROC 0.66). Additional research is needed with larger sample sizes to validate our initially promising findings.

Results from prior studies indicated that the use of information about the history of previous respiratory states may be a promising approach for predicting the duration of apneic events. A study of capnography waveform abnormalities during nurse-administered sedation found a two-fold increase in the risk of apnea (hazard ratio [HR] 2.14; 95% CI 1.75-2.62) when a patient was in a state of hypoventilation (defined as >10% reduction in end-tidal CO₂ from baseline) [24]. The risk of apnea also increased with each additional sedative dose (HR 2.86; 95% CI 2.15-3.81) [24]. Results from an earlier study in a different population also supported the observations that the onset of apneic periods during sedation is associated with a previous history of abnormal respiratory state. Krauss and colleagues [25] used survival analysis to model the time to first apneic events in a sample of 312 patients undergoing procedural sedation with propofol or ketamine in the emergency department. They found that the risk of apnea increased with an abnormal end-tidal CO₂ measurement 30 seconds (HR 2.45; 95% CI 1.63-3.69), 60 seconds (HR 1.88; 95% CI 1.21-2.92), and 90 seconds (HR 2.06; 95% CI 1.36-3.11) prior to an apneic event. In our study, we leveraged information about the associations between apneic events and the history of previous respiratory states by building a predictive model using a machine learning approach. Features included in the models we tested were the previous respiratory state, duration of time in the previous respiratory state, number of previous apneic events, and duration of the previous apneic event.

Discussion

In this study, we found that the random forest model had the best discriminative ability and calibration for predicting if an apneic event would be prolonged during nurse-administered procedural sedation. However, it should be noted that the accuracy of this random forest model was still quite low (AUROC 0.66). Additional research is needed with larger sample sizes to validate our initially promising findings.

Results from prior studies indicated that the use of information about the history of previous respiratory states may be a promising approach for predicting the duration of apneic events. A study of capnography waveform abnormalities during nurse-administered sedation found a two-fold increase in the risk of apnea (hazard ratio [HR] 2.14; 95% CI 1.75-2.62) when a patient was in a state of hypoventilation (defined as >10% reduction in end-tidal CO₂ from baseline) [24]. The risk of apnea also increased with each additional sedative dose (HR 2.86; 95% CI 2.15-3.81) [24]. Results from an earlier study in a different population also supported the observations that the onset of apneic periods during sedation is associated with a previous history of abnormal respiratory state. Krauss and colleagues [25] used survival analysis to model the time to first apneic events in a sample of 312 patients undergoing procedural sedation with propofol or ketamine in the emergency department. They found that the risk of apnea increased with an abnormal end-tidal CO₂ measurement 30 seconds (HR 2.45; 95% CI 1.63-3.69), 60 seconds (HR 1.88; 95% CI 1.21-2.92), and 90 seconds (HR 2.06; 95% CI 1.36-3.11) prior to an apneic event. In our study, we leveraged information about the associations between apneic events and the history of previous respiratory states by building a predictive model using a machine learning approach. Features included in the models we tested were the previous respiratory state, duration of time in the previous respiratory state, number of previous apneic events, and duration of the previous apneic event.
Many prediction modeling studies focused on predicting clinical outcomes have yielded similarly low AUROC scores. For example, a recent study of the predictive ability of vital sign parameters for clinical deterioration in subacute care patients reported an AUROC score of 0.57 [26]. Decision curve analysis can help elucidate whether a model with low AUROC scores is “good enough” to use in practice. Our results indicated that nurses currently using the conservative strategy who are willing to value a false positive about 2-3 times more than a true positive would not derive an overall net benefit from using the random forest model as a trigger for apnea alarms. This is because the random forest model would produce a worse outcome than the default strategy of waiting for an alarm to be triggered at the 30-second threshold in terms of the balance between true positives and false positives for determining if an apneic event will be prolonged. Conversely, nurses currently using the aggressive strategy who are willing to value a false positive about 2-3 times more than a true positive would derive an overall net benefit from using the random forest model as a trigger for apnea alarms. Using the random forest model as an additional input for an alarm trigger would reduce the total alarm burden and could be considered an option for implementation into practice. To operationalize these predictions into capnography monitors, partnerships with industry would be required because monitor functionality would need to be adapted to facilitate input of the data required to calculate the predictions [27]. These data would include patient characteristics and sedative dosing. Integrating predictive models into alarm management strategies for respiratory monitoring devices is also indicated in other contexts. For example, a recent study found that opioid-induced respiratory depression during recovery from anesthesia can be accurately predicted using a machine learning approach [28].

In addition, user-centered design considerations, such as how the predictions should be communicated to nurses responsible for decision-making, are important avenues for further research prior to implementation [29]. This study used decision curve analysis to evaluate the potential clinical impact that using the model as input for capnography alarm management would have on the number of alarms triggered (ie, false positives and false negatives). However, as with any intervention in health care, the efficacy of the model needs to be assessed prior to broader implementation. The indicator for efficacy in this context would be the improvement in patient safety using this model as input for capnography alarm management. The gold standard approach for such an evaluation is a randomized controlled trial. Randomized controlled trials testing alarm conditions that have integrated predictions from machine learning models have been conducted previously in similar contexts such as intraoperative blood pressure management [30,31].

A noteworthy finding is that the model produced an overall net benefit that was higher than that of the aggressive strategy but not higher than that of the conservative strategy. Further research with larger sample sizes is needed to increase the predictive power of models aimed at predicting the duration of apneic events. Such research is warranted because triggering an alarm after 30 seconds of apnea that would turn off without clinical intervention only 5 seconds later is just as clinically inconsequential as triggering an alert after 5 seconds of apnea that would similarly turn off after a short time. In both these circumstances, there would not be enough time for the clinician’s intervention to take effect. However, presumably in an attempt to reduce alarm burden, the default settings for many capnography monitoring devices are for the alarm to be triggered after 30 seconds of apnea. An ideal alternative to the conservative strategy would be for capnography monitor alarms to be triggered as early as possible during an apneic event, but only if the event will be prolonged enough to necessitate clinical intervention and for this intervention to take effect—a goal that we did not achieve in this study. However, previous research indicates that it would be worthwhile to find such a solution. An analysis of half a million patients found that respiratory compromise during interventional radiology procedures performed with moderate sedation led to worse clinical outcomes and higher costs than those observed in normal respiratory states [32].

Limitations

Although the number of apneic events included in the models was relatively high, this was seen in a small number of patients. We used cross-validation to minimize the possibility of overfitting. This analysis used data from an observational study conducted at two hospitals that used a convenience sampling approach; therefore, selection bias was possible. The context in which the study was conducted should also be considered in terms of external validity. Participants were patients undergoing procedures in a cardiac catheterization laboratory where small bolus doses of midazolam and fentanyl were used for sedation. Other procedural sedation contexts may involve the use of different sedative doses and types of medications, so the results of this study might not be generalizable to these contexts. A further limitation is that clinicians were not blinded to capnography measurements because of the observational nature of the study design. It is possible that interventions used by clinicians during the 0- to 30-second apneic period influenced the duration of the apneic event. However, this mimics real-world practice in that interventions may be implemented at clinicians’ discretion where no alarm conditions have been met. Additionally, 25% of the study population had sleep apnea, which was one of the predictors included in the model. Due to the small sample size, the dataset used to train the model would have contained only a small proportion of patients with sleep apnea and therefore it may not be generalizable to the larger population of individuals with sleep apnea. Further research with larger sample sizes is required to confirm our findings.

Conclusion

We evaluated several candidate models to determine their accuracy in predicting at the 15-second timepoint if an apneic event would prolong for >30 seconds. The random forest model performed the best in terms of discrimination and calibration. Decision curve analysis indicated that using the random forest model for capnography alarm management would lead to a better outcome than using an aggressive strategy in which alarms are triggered after 15 seconds of apnea. The model would not
be superior to the conservative strategy in which alarms are only triggered after 30 seconds.

Acknowledgments
AC received support from a National Health and Medical Research Council early career fellowship (APP1091657) and the Bertha Rosenstadt Faculty Small Research Grant from the Lawrence S. Bloomberg Faculty of Nursing at the University of Toronto. This study received funding from Wesley Medical Research (project number: 2015-31). The funders had no role in the study design, in the collection, analysis, and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

Conflicts of Interest
None declared.

References


Abbreviations

- AUROC: area under the receiver operating characteristic curve
- CO2: carbon dioxide
- HR: hazard ratio