JMIR Perioperative Medicine

Technologies for pre- and post-operative education, preventative interventions, and clinical care for surgery and anaesthesiology patients, as well as informatics applications in anesthesia, surgery, critical care, and pain medicine
Volume 9 (2026) ISSN 2561-9128 Editor in Chief: Nidhi Rohatgi, MS, MD, SFHM

Contents

Review

Original Papers

Multidimensional Assessment of Recovery After Total Knee Arthroplasty in Clinical Practice: Critical Narrative Review

Abderrahmane Boukabache[1], BSc (Hons), MSc; Nimalan Maruthainar[2], BSc (Hons), MBBS; Vikrant Manhas[3], MBBS, MS; Darren Player[1], BSc (Hons), PhD

[1]Division of Surgery and Interventional Science, Faculty of Medical Sciences, University College London, Bloomsbury Campus: 2nd Floor, Charles Bell House, 43-45 Foley Street, London, United Kingdom

[2]Department of Trauma and Orthopaedics, Royal Free Hospital, London, United Kingdom

[3]Department of Orthopaedics, All India Institute of Medical Sciences, New Delhi, India

Corresponding Author:
Abderrahmane Boukabache, BSc (Hons), MSc
Division of Surgery and Interventional Science, Faculty of Medical Sciences, University College London, Bloomsbury Campus: 2nd Floor, Charles Bell House, 43-45 Foley Street, London, United Kingdom

Abstract

Background: Total knee arthroplasty (TKA) is the primary treatment for advanced knee osteoarthritis. Despite its clinical success and favorable patient-reported outcome measures (PROMs), approximately 20% to 30% of patients continue to experience persistent functional limitations and muscle weakness. This highlights the need for a comprehensive evaluation of recovery parameters beyond pain and range of motion. Given the wide range of methods available for assessing TKA outcomes, clinicians often select tools based on personal preference and understanding, which may affect accuracy and consistency; for example, the Knee Injury and Osteoarthritis Outcome Score may overestimate function compared to gait analysis studies.

Objective: The aim of this study was to conduct a narrative review focusing on the use, strengths, and limitations of different outcome measures used in routine orthopedic practice to optimize post-TKA evaluation.

Methods: A literature search was conducted in February 2025 across 2 databases (PubMed and Web of Science). Eligible studies included original research articles, systematic reviews, and meta-analyses that focused on validated measures used to evaluate TKA. Case reports, conference abstracts, and studies focused exclusively on surgical techniques were excluded. Themes were identified across studies to structure the results according to types of assessments and clinical applicability.

Results: A total of 6831 studies were retrieved and screened in this review, with 4 themes emerging around muscle mass, strength, performance, and PROMs. The Oxford Knee Score is favored for its ease of use and minimal ceiling effects. Broader tools like the Knee Injury and Osteoarthritis Outcome Score and Western Ontario and McMaster Universities Osteoarthritis Index provide detailed insights but are less practical clinically. For muscle strength, the portable fixed dynamometer showed high reliability and comparability to isokinetic dynamometry. Dual-energy X-ray absorptiometry remains the gold standard for assessing muscle mass, while bioelectrical impedance analysis offers a practical alternative. The 5-Repetition Sit-to-Stand test effectively evaluates lower limb power and speed.

Conclusions: Clinicians should integrate both objective (muscle mass, strength, and performance) and subjective (PROMs) measures to improve TKA recovery assessment. This multidimensional approach has the potential to enhance the accuracy of patient evaluation and supports the development of tailored rehabilitation strategies that address individual deficits and optimize functional outcomes.

(JMIR Perioper Med 2026;9:e84011) doi:10.2196/84011

KEYWORDS

arthroplasty; osteoarthritis; function; outcome; total knee replacement

Introduction

Osteoarthritis is a leading cause of pain and disability worldwide, affecting 528 million people as of 2019 [1,2]. Among various treatment options for osteoarthritis, total knee arthroplasty (TKA) stands out for its substantial impact on alleviating chronic knee pain [3]. With advancements in prosthetic design and surgical techniques, TKA has demonstrated high survival rates and relatively low complication risks, making it one of the most successful orthopedic procedures. However, despite these clinical successes, postoperative functional outcomes can vary among patients [4], with approximately 20% to 30% experiencing persistent functional limitations and muscle weakness [5].

Several factors influence recovery following TKA, including surgeon-related aspects such as surgical volume and technique, as well as patient-related variables like preoperative physical conditioning and psychological status. Among these, preoperative quadriceps strength has been identified as a key predictor of postoperative function, directly impacting mobility and the ability to perform daily activities [6]. Nonetheless, frequently used tools to evaluate TKA in orthopedics often rely on measures such as pain assessment and range of motion (ROM) [7], which, while valuable, may not fully capture postoperative recovery or functional capacity.

Evidence suggests that these metrics correlate poorly with objective measures of physical function and do not adequately reflect muscle weakness or biomechanical impairments that persist post-TKA [8]. While patient-related outcome measures (PROMs) provide subjective insights into perceived recovery, they may overestimate functional improvements compared to more objective assessments, such as gait analysis and performance-based tests [9]. This discrepancy underscores the need for a more comprehensive evaluation framework that integrates multiple dimensions of recovery, including muscle mass, muscle strength, and physical performance.

This review aims to critically examine the use, strengths, and limitations of different outcome measures used in routine practice when assessing recovery post-TKA, emphasizing the importance of incorporating objective physical performance metrics alongside PROMs. We hypothesize that a multidimensional approach, combining assessments of muscle mass, strength, physical performance, and PROMs, will provide a more accurate and clinically meaningful evaluation of TKA recovery, ultimately guiding more effective rehabilitation strategies.

Methods

A literature search of studies focusing on methods assessing TKA outcomes was performed using PubMed and Web of Science databases in February 2025. No limits were applied to publication dates. Inclusion criteria comprised original research, systematic reviews, and meta-analyses. After the removal of duplicates and cross-referencing of articles, we screened titles and abstracts for relevance. Studies were excluded if they were case reports, conference abstracts, or focused exclusively on surgical techniques. Only studies using standardized and validated assessment tools were included to ensure clinical relevance. The methodological quality of included studies was evaluated based on study design, sample size, and clarity of outcome measures.

One author conducted the initial screening of all retrieved abstracts against the inclusion criteria. Full texts of potentially relevant studies were then independently reviewed for eligibility and data extraction. Discrepancies between reviewers were resolved through discussion until consensus was achieved. Extracted data were thematically organized (muscle mass, muscle strength, physical performance, and PROMs), allowing patterns across assessment domains to be identified and synthesized into a structured narrative aligned with the review's objectives.

Given the narrative nature of this review, our selection was not strictly systematic; instead, we documented the total number of retrieved and screened articles (n=6831).

Results

Patient Reported Outcome Measures

There are numerous PROMs available to assess the outcomes of TKA, including: Knee Society Clinical Rating System (KSS), Western Ontario and McMaster Universities Osteoarthritis (WOMAC), Knee Injury and Osteoarthritis Outcome Score (KOOS), and the Oxford Knee Score (OKS) [10-13]. In clinical practice and research, it is essential to prioritize the use of sound PROMs over frequently used ones. The ultimate goal herein is to find a balance between standardized measures and specific contextually relevant PROMs in the field of TKA.

The KSS and OKS PROMs are shorter than the KOOS and WOMAC, with 10 and 12 items, respectively. The OKS uses a 4-week recall period, while the patient-reported component of the KSS is designed to reflect current knee status without a fixed recall window, which may reduce recall bias but limit comparability across time points. The OKS is solely derived from patient input, while the KSS comprises a PROM section completed by the patient and an informational-clinical section for the surgeon. Only the PROM section of the KSS is used to produce a psychometrically valid knee score. This separation sought to allow the functional PROM section to be independently assessed from the clinical section and confounding factors such as age or comorbidities. Notably, Martimbianco et al [14] reported considerable inter- and intraexaminer variability in the clinical-reported section of the KSS, which raises concerns about its reliability and consistency in clinical use. In contrast, the KOOS and WOMAC are larger instruments, with 42 and 33 items, respectively. They assess symptoms, stiffness, pain, and activities of daily living (ADL), but also include additional domains such as sports or recreation and quality of life. Like the OKS, these instruments are totally focused on the patient and their self-reported experiences. The WOMAC uses a 48-hour recall period, which minimizes recall bias and improves measurement precision; however, this narrow timeframe may fail to reflect symptom variability and functional fluctuation over longer recovery periods following TKA. In contrast, the KOOS uses a 1-week recall period, providing a broader representation of knee symptoms and functional limitations. Nevertheless, this longer recall window may increase susceptibility to recall bias and potentially reduce sensitivity to short-term clinical change.

All 4 PROMs assess pain and activity function, but they differ in how they capture pain information. The OKS and KSS both include a single item that captures the patient's general level of pain. However, the OKS asks further pain information in the context of general activities, such as work interference and walking. This is particularly important to assess as movement is associated with changes in pain, from subtle changes in muscle coordination to complete avoidance of joint function [15]. In contrast, the KOOS and WOMAC both include a section dedicated to assessing pain experienced during specific activities, with nine identical items. This encourages the patient to recall and report on the pain experience during knee maneuvers such as knee twisting and straightening. The KOOS is an extension of the WOMAC, which justifies the similarities in the items. It is used in younger and/or more active patients who typically perform demanding twisting movements in their ADL. By assessing knee pain during movements like twisting and straightening, changes in functional outcomes over time can be tracked. This is especially relevant post-TKA, where pain during specific activities is directly linked to the functionality of the joint and the surrounding muscles [16].

The only resemblance when assessing function across the 4 PROMs is the patient's comfort level when handling stair ascent and descent. This likely reflects the universal importance of these activities in daily life and the significant impact that knee function has on an individual's mobility and well-being. However, each PROM also has unique items that assess different aspects of knee function and quality of life.

The OKS, KOOS, and WOMAC all assess the knee concerning various ADL, such as kneeling, transportation, domestic work, and bathroom activities. The OKS and KSS share only the item related to the distance the patient can walk. The KSS and items within the symptoms section of both KOOS and WOMAC ask about the patient's ability to fully bend and extend the knee, as well as objective measures of flexion contracture and extension lag. The use of similar items and domains across these scores can lead to greater consistency in assessing knee function and symptoms. Furthermore, consistency is important for tracking changes over time and comparing outcomes across different patients and studies. The KOOS is unique in that it assesses ADL related to sports and recreation and includes items that ask about quality of life. While similarities can provide consistency, differences between PROMs may capture important nuances specific to each condition.

The KSS evaluates 4 distinct domains: clinical, functional, satisfaction, and fulfillment of expectation. A notable feature is the inclusion of both high-demand tasks and 3 patient-selected priority activities from a predefined list, aiming to tailor the assessment to individual goals and enhance the relevance of functional evaluation. This patient-led component is designed to offer a more individualized perspective on recovery and may help inform more targeted postoperative rehabilitation. However, despite these strengths, the validity of the KSS has been questioned, particularly in the context of TKA outcome measurement. Several studies have noted limitations, including Ghanem et al [17], who highlighted concerns in revision TKA scenarios, and Vogel et al [18], who found the system less responsive than broader tools like the 36-Item Short Form Survey and WOMAC. One key criticism lies in the item-selection process: patients were not involved in the development of the tool, which may result in content that does not fully reflect patient priorities or lived experiences. Bach et al [19] further noted that the relatively limited item pool may constrain the scale's ability to capture the full spectrum of functional outcomes. This lack of patient input and limited item scope could contribute to a misalignment between what the tool measures and what matters most to patients, ultimately impacting its content validity and clinical use [17,18,20].

Ghanem et al [17] highlight that 2 very different patients could receive the same clinical score on the KSS. For instance, a patient with a stiff, pain-free, well-aligned knee and a patient with mild pain, excellent knee motion, and normal alignment may both receive a similar score, even though they have very different levels of function in their daily lives. To address these concerns, a revised Knee Society Scoring System (2011 Knee Society Score) was developed and validated as a more reliable measure for assessing outcomes in TKA procedures [21,22]. Albeit, a revised outcome measure, PROMs have inherent limitations in that they rely primarily on self-reported data, which may not always align perfectly with objective measures of physical function. This is where muscle function analysis could come into play. It offers an objective assessment of a patient's physical capabilities, including muscle strength, range of motion, and muscle activation patterns. By quantifying these aspects, health care providers can validate and complement the patient's self-reported symptoms and limitations with tangible data. This objective information not only supports the patient's reported experiences but also provides a more complete picture of their health status.

Despite its validity issues, the KSS is still popular among clinical researchers [23], perhaps for including alignment and ROM measurement. Proper coronal alignment of the components in TKA has been shown in the literature to be critical for implant survival rate. Additionally, knee ROM is a crucial indicator of successful TKA and is necessary for many ADLs [24]. While ROM is commonly measured and reported, the significance of isolated muscle function has been overlooked here. Capin et al [25] found quadriceps strength in particular to be a critical factor in TKA recovery and considered a rate-limiting step. This limited quantification of isolated muscle function is an area that warrants attention and improvement in the care of patients who underwent TKA.

The OKS was developed specifically for evaluating the outcomes of TKA. The simplicity and shortness of the questionnaire make it an attractive option for clinicians, and this may in part have contributed to its broad use in cohort studies and joint replacement registers [26,27]. This intentional oversimplification of the questionnaire highlights a lack of scope and sophistication to adequately capture the complex, interrelated issues experienced by many patients, such as joint stiffness, muscle weakness, and instability. To address these limitations, the use of a combination of assessment tools and methods has been suggested [28]. This could include performance-based functional tests, isolated muscle function tests, muscle mass measurement, and imaging studies.

Another consideration is that OKS has a high response rate when compared to other PROMs [10]. Analysis of the National Health Service PROMs dataset of 72,154 OKS concluded that OKS does not exhibit a ceiling or a floor effect at 6 months [29]. Marx et al [30] reported a ceiling effect of 22% at 12 months following surgery, but this could be attributed to patients achieving an optimal outcome rather than a limitation of the OKS. Nevertheless, this could indicate that the OKS may be more appropriate for short-term outcomes and might inadequately reflect the long-term burden. Although it has been used in randomized controlled trials to assess knee arthroplasty long-term outcomes [31].

The WOMAC, developed in 1982, has undergone multiple revisions since and has been validated for TKA clinical trials [32-34], with ceiling effects at 6 months and 12 months for patients who underwent TKA [35]. The KOOS was an upgrade to the WOMAC to effectively capture the requirements of younger and more active individuals with knee injury or osteoarthritis. Several studies have shown that the KOOS is more sensitive and responsive than the WOMAC in younger or more active patients with knee injury or knee osteoarthritis [36]. Roos and Toksvig-Larsen [35] evaluated the outcome of 105 patients (mean age of 71) after TKA and found 74% of all Sport/Recreation items were considered to be “not applicable.” The floor effects of approximately 48% likely reflect the high-demand activities, such as sports and recreation, which may be more relevant for younger, more active adults undergoing TKA.

A systematic review of the literature performed by Collins et al [37] found, as it was intended, the KOOS’s ADL subscale has better content validity for older patients, and Sport/Recreation has better content validity for younger patients. Increasing ceiling effects from 6 to 12 months, particularly in the pain domain, may suggest that some patients have reached a plateau in their recovery, but also raises questions about the sensitivity of the measurement tool.

According to the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN), which define internationally accepted criteria for evaluating the validity, reliability, and responsiveness of PROMs, the OKS meets COSMIN requirements and is recommended for use as a TKA outcome measure. Pain and function subscales of WOMAC and KOOS also demonstrate adequate measurement properties when evaluated as standalone subscales. However, the KSS does not consistently fulfill all COSMIN criteria and is not classified among the instruments meeting minimum standards for psychometric validation [38]. Several studies have evaluated the performance of different PROMs in measuring outcomes following lower extremity joint replacement surgery. Harris et al [39] identified the OKS and WOMAC as the best-performing PROMs specific to the lower extremity. The study assessed the validity, reproducibility, and acceptability of the scoring systems. Similarly, Alviar et al [40] in their review found OKS and WOMAC to be the best PROMs after assessing the quality of patient-reported outcome scoring systems. Collins and Roos [41] analyzed attributes of the 11 most frequently used PROMs for total hip replacement and TKA and considered KOOS, WOMAC, and OKS to be good PROMs. These PROMs aid in monitoring progress, guiding interventions, and facilitating shared decision-making. Nevertheless, there are limitations, such as the risk of subjective bias and variability in interpretation. Patient responses may be influenced by personal perceptions, mood, cognitive state, or cultural differences. Obtaining consistent and accurate data can be challenging, especially if patients struggle to recall specific details over time. To gain a more thorough understanding of the impact of TKA, it is crucial to integrate objective measures of muscle strength and functional assessments, particularly in addressing muscle weakness.

Strength Measurements

The indirect assessment of muscle function in any of the PROMs may not provide sufficient indication of dysfunction, necessitating a comprehensive analysis of objective assessment tools in TKA (Table 1)(Multimedia Appendix 1). To this end, this section will focus on the types of muscle strength assessments that are feasible in a TKA clinical setting.

Table . Comparative analysis of outcome measures used in total knee arthroplasty (TKA), including patient-reported outcomes, muscle strength, muscle mass, and physical performance tools. Comparison criteria include measurement type, reliability, validity, ease of use, clinical relevance, limitations, and best use case.

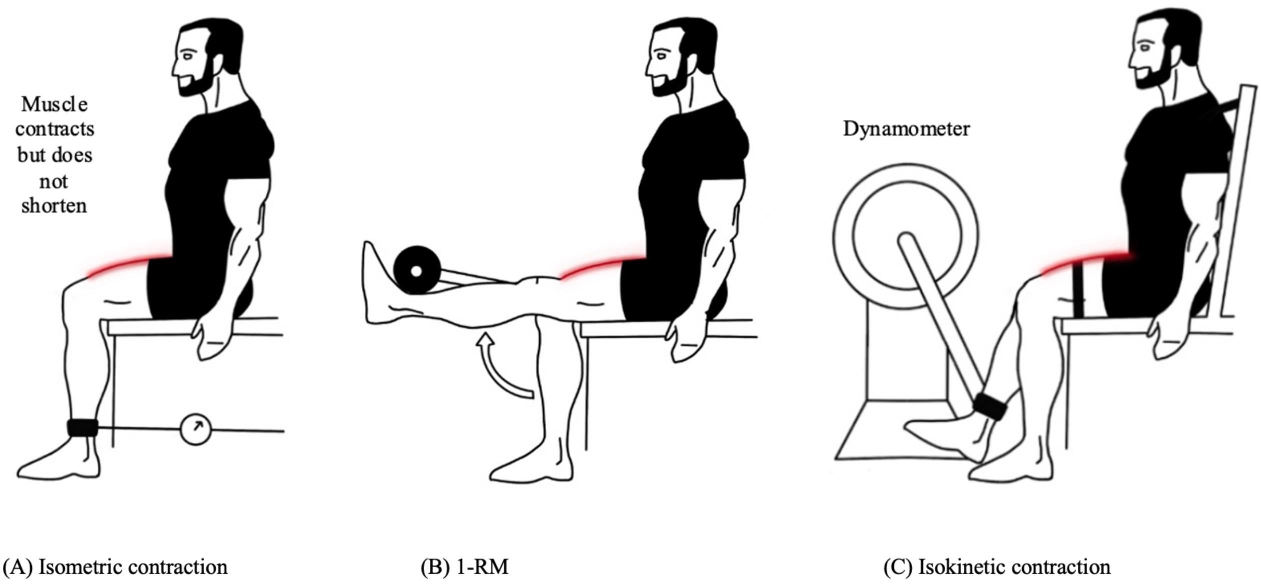
Comparison	PROMs[a] (KSS[b], WOMAC[c], KOOS[d], and OKS)[e]	Muscle strength tests (IKD[f], PFD[g], and HHD[h])	Muscle mass tests (CT[i], MRI[j], DXA[k], and BIA[l])	Physical performance tests (TUG[m], 6MWT[n], 5R-STS[o], and SCT[p])
Measurement type	Subjective (patient-reported pain, function)	Objective (muscle force output)	Objective (muscle cross-sectional area or mass)	Objective (functional movement performance)
Reliability	High test-retest reliability but can be influenced by patient perception	High reliability but dependent on test consistency and patient effort	High for CT, MRI, and DXA; moderate for BIA (affected by hydration)	High for TUG, 6MWT, and 5R-STS
Validity	Valid for subjective function but may not correlate with actual muscle strength or mobility	IKD, highest validity. PFD, good validity with proper fixation. HHD, lower validity	Highly valid for assessing atrophy or hypertrophy	Strong correlation with mobility and functional independence
Ease of use	Simple, noninvasive, and patient-friendly	Requires equipment; portable dynamometers are easier than isokinetic ones	CT or MRI is expensive; DXA is more accessible; BIA is easiest	Quick, requires minimal equipment (eg, chair and stopwatch)
Clinical relevance	Useful for tracking patient-perceived recovery	Directly assesses quadriceps and hamstring strength, critical for TKA recovery	Helps detect muscle loss post-TKA, which can affect long-term function	Strong predictor of fall risk, mobility, and independence
Limitations	Subjective, may not reflect real functional capacity	Requires patient cooperation; costly for isokinetic dynamometers	Expensive (CT and MRI), radiation exposure (CT and DXA), less precise (BIA)	Can be influenced by patient motivation, fatigue, or comorbidities
Best use case	Tracking patient-perceived progress	Measuring post-TKA quadriceps or hamstring recovery	Evaluating long-term muscle loss and sarcopenia	Assessing mobility and real-world function

[a]PROM: patient-reported outcome measure.
[b]KSS: Knee Society Score.
[c]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.
[d]KOOS: Knee Injury and Osteoarthritis Outcome Score.
[e]OKS: Oxford Knee Score.
[f]IKD: isokinetic dynamometer.
[g]PFD: portable fixed dynamometer.
[h]HHD: handheld dynamometry.
[i]CT: computed tomography.
[j]MRI: magnetic resonance imaging.
[k]DXA: dual-energy X-ray absorptiometry.
[l]iBIA: bioelectrical impedance analysis.
[m]TUG: Time Up and Go.
[n]6MWT: 6-Minute Walk Test.
[o]5R-STS: 5-Repetition Sit-to-Stand.
[p]SCT: Stair Climb Test.

When evaluating muscle function, the measured outcomes commonly include muscle strength and power [42]. Various methodologies assess these parameters, ranging from manual muscle testing to more complex and expensive isokinetic assessments. However, no general consensus exists regarding the preferred method of assessing muscle function in orthopedic practice, as each approach has distinct advantages and limitations [43].

Isometric strength assessment measures the maximum force generated during a contraction in which the muscle length remains constant (Figure 1A). Although isometric contractions are relatively uncommon in daily activities, their clinical relevance lies in their ability to predict functional capacity, particularly in older individuals and those with significant functional impairments [44]. Furthermore, isometric strength assessments have demonstrated high reliability across different orthopedic populations [45,46]. Isometric strength assessments can be easily performed using a dynamometer or force plate, and the individual is instructed to push or pull against an immovable object. Portable fixed dynamometer (PFD) has demonstrated high reliability (intraclass correlation coefficient [ICC] >0.90) and validity in measuring strong muscle contractions such as knee extension [47-49]. Importantly, studies have shown that knee muscle strength measurements using a PFD highly correlate with those obtained from an IKD, making PFD a viable alternative for clinical assessment [50].

Figure 1. Illustration of 3 common muscle strength assessment methods used in total knee arthroplasty (TKA) recovery: (A) isometric contraction—muscle generates force without joint movement; (B) one-repetition maximum (1-RM)—maximal weight lifted once with proper form; and (C) isokinetic contraction—muscle contracts at a constant speed throughout the range of motion using a dynamometer.



Toonstra and Mattacola [51] compared the reliability of IKD, PFD, and handheld dynamometry for isometric knee strength assessment. Their findings revealed high test-retest reliability for both IKD and PFD, while handheld dynamometry showed fair to poor reliability.

The strength of IKD lies in its ability to provide stabilization during testing and standardized protocols, making it the gold standard. However, PFD demonstrated comparable reliability while offering greater portability, ease of use, and cost-effectiveness. This makes PFD an attractive option for routine clinical evaluations, particularly in settings where IKD is not feasible.

Another key advantage of isometric assessments is their safety. Unlike dynamic strength tests, which involve movement and may place stress on healing tissues, isometric assessments minimize joint strain. This makes them particularly suitable for early postoperative assessments, where patient safety is a priority. Given these benefits, isometric strength measurements serve as a practical and reliable method for monitoring recovery following TKA.

One-repetition maximum (1-RM): Muscle strength can also be assessed through the one–repetition maximum (1-RM) test, which determines the maximum load that can be lifted in a single attempt, such as during a squat or leg extension (Figure 1B). This method evaluates muscle strength in a more functional and dynamic context, offering insight into the ability to generate force during real-world movements.

Despite its functional relevance, 1-RM testing has limitations. The requirement for specialized equipment (eg, weights and resistance machines) and the time-intensive nature of testing make it less practical for large patient cohorts. Additionally, 1-RM testing carries a risk of injury, particularly in postoperative populations where patients may not yet be able to safely perform maximal-effort movements. To mitigate this risk, researchers often use submaximal testing protocols to estimate 1-RM values [52].

Isokinetic testing measures muscle strength as the peak torque generated during a contraction performed at a constant angular velocity (Figure 1C). This method provides resistance that matches the individual's effort throughout the ROM, allowing for maximal force production at different joint angles. Isokinetic assessments are particularly valuable for evaluating muscle imbalances, monitoring rehabilitation progress, and understanding strength deficits following TKA.

Despite these advantages, isokinetic testing has notable limitations. The high cost and lack of portability make it less accessible in many clinical settings [53,54]. Additionally, isokinetic assessments require specialized equipment and trained personnel, making them impractical for routine postoperative evaluations. The potential for joint irritation and discomfort during testing further limits its suitability for early rehabilitation stages.

A key consideration when choosing between isometric and isokinetic testing is clinical feasibility. While isokinetic assessments provide detailed torque-angle relationships, they are not always necessary for functional recovery monitoring. In contrast, isometric testing provides a simple, reliable, and cost-effective means of assessing muscle strength, making it better suited for routine evaluations.

Lauermann et al [55] compared quadriceps strength assessments using isometric, isokinetic, and 1-RM methods in patients who underwent TKA. All 3 methods showed similar validity in detecting strength deficits, suggesting that isometric testing is an effective alternative to more complex assessments. Similarly, Lienhard et al [45] found comparable test-retest reliability across these methods in patients who underwent TKA, reinforcing the viability of isometric assessments in clinical practice.

Muscle Mass Measurements

Studies have shown a significant correlation between skeletal muscle mass and cross-sectional skeletal muscle area in the extremities with muscle strength or power [56,57]. Hence, assessing these morphological features could be a good way to gauge muscle function. Furthermore, with certain imaging modalities, it may also be possible to determine the extent of sarcopenia, muscle wasting, and the level of fibrosis and fat infiltration—all factors that play a significant role in muscle function.

Computed tomography (CT) is an imaging modality that was introduced over 50 years ago, as the first clinically accepted tool for body composition measurement and served as a gold standard [58]. This method uses X-ray beams to create cross-sectional images of the body, allowing estimation of total body fat, visceral fat, and skeletal muscle mass [59,60]. Despite its early acceptance and high validity for assessing limb muscle cross-sectional area, with excellent test–retest and inter-observer reliability (ICC 0.98 - 1.00 for thigh muscle measurements), CT has general limitations such as high costs, the need for skilled technicians, and radiation exposure [61,62]. However, in current practice, CT is rarely used for measuring muscle mass, likely due to its associated limitations and the emergence of alternative, more practical techniques such as magnetic resonance imaging (MRI).

The introduction of MRI in the 1980s replaced the CT scan as the gold standard [63]. Its 3D images of skeletal muscle, fat tissue, and other organs are created by emitting radiofrequency energy from hydrogen atom nuclei in magnetic fields, with signal variations indicating different tissue types [63,64]. This development brought about high-resolution images without radiation exposure, making it suitable for tracking small changes over time, which is beneficial for intervention and observational studies. Advancements in MRI techniques have notably reduced image acquisition times, and modern scanners can accommodate obese individuals. Validated against direct anatomical measurements (r=0.97), MRI exhibits high test-retest and inter-observer reliability (2.9%, r=0.99% and 2.6%, r=0.99) in healthy populations [63,64]. However, limitations in clinical and research settings include high costs, the need for technical expertise, space requirements, infeasibility for patients with claustrophobia or those with MRI-incompatible implanted devices (eg, cardiac pacemakers), and standardization is hampered by the existence of various data acquisition protocols [61,65].

Dual-energy X-ray absorptiometry (DXA) has emerged as a strong alternative to the gold standard for assessing body composition, being relatively cheap compared with CT scan and MRI, and easy to perform. Initially designed for measuring bone mineral density, DXA is now widely used for examining overall body composition and muscle mass [63]. DXA operates by using 2 X-ray beams to differentiate between fat, bone, and lean tissue based on their X-ray absorption properties [66]. This technique has shown a high correlation with both MRI and CT in estimating skeletal muscle mass, indicating its reliability (r=0.88 and r=0.77 - 0.95, respectively) [67-69]. Additionally, DXA test-retest reliability for measures of fat-free mass demonstrated a high correlation (r=0.99), with low precision errors ranging from below 1% to 3% [70].

However, DXA does present some limitations. First, it does not directly quantify muscle mass; instead, it calculates lean soft tissue mass or fat-free mass based on the different gray tones observed in the DXA scan. Assumptions are made during this calculation, and factors like dehydration and edema, which can be common in populations with obesity, may interfere with the accuracy of the measurements. Additionally, the lack of standardization across devices, software packages, and versions can yield different results, affecting the reliability and comparability of measurements. Moreover, while DXA involves less radiation exposure compared to CT, there is still some exposure [71,72].

Bioelectrical impedance analysis (BIA) has become widely used for assessing body composition in both clinical and nonclinical settings [73]. BIA operates on the principle that tissues with higher water and electrolyte content, like skeletal muscle, offer less resistance to the passage of a low-voltage electrical current compared to lipid-rich adipose tissue, such as bone. This conductivity difference is exploited by BIA systems to quantify different body compartments. Measurements obtained along with parameters like sex, age, and body weight are then integrated into the population-specific body composition prediction equations [64,72]. The advantages of BIA stem from its noninvasive nature, cost-effectiveness, and ease of use, but limitations include sensitivity to factors like body position, physical exercise, food intake, hydration status, and the need for population-specific equations [74]. A study by Buckinx et al [75] attempted to gauge the reliability of BIA in assessing appendicular lean mass. The results showed high intraoperator reliability, indicated by an ICC of 0.89 when performed by the same operator, and interoperator reliability was also relatively high with an ICC of 0.77 when performed by different operators. However, the study revealed a notable discrepancy when comparing appendicular lean mass assessed by DXA and predicted by BIA, as evidenced by a low ICC of 0.37. Importantly, there exists a potential for a significant prediction error at the individual level with BIA, coupled with a systematic positive bias leading to an overall underestimation of lean body mass measurements [76]. Additionally, hydration status is a frequent confounder in clinical settings, as BIA relies on electrical conductivity that is highly sensitive to fluctuations in body water content, potentially compromising accuracy and limiting its reliability in certain patient populations. Despite these limitations, BIA remains a viable alternative in situations where more precise methods are not feasible.

Performance Measurements

While measurements of muscle mass and function are important for understanding physiological dysfunction and pathology, assessing physical performance is essential. In this context, functional performance-based tests use an individual’s body weight as resistance, quantifying performance by the time taken or the number of repetitions completed. Several authoritative groups, including The Osteoarthritis Research Society

International, Rehabilitative Care Alliance, and the American Physical Therapy Association, recommend performance-based measures for individuals with osteoarthritis and arthroplasty [77-79]. These measures typically include assessments of gait speed tests, Stair Climb Tests (SCT), and Sit-to-Stand (STS) tests, each with distinct yet overlapping focuses.

Gait speed tests primarily assess mobility, endurance, and functional independence by measuring the time required to walk a set distance. They are simple, reliable, and highly predictive of overall recovery, fall risk, and mortality [80]. SCTs focus on functional strength, coordination, and balance, as stair negotiation requires greater knee flexion and quadriceps activation than level walking [81]. SCTs are more sensitive to persistent functional deficits than gait speed tests, as stair climbing is often more challenging postoperatively. STS tests, on the other hand, primarily evaluate lower limb strength and endurance by measuring the ability to transition from sitting to standing multiple times within a set duration or repetition count. STS tests strongly correlate with quadriceps strength and are useful for monitoring functional recovery [82].

While all 3 measures provide valuable insights into TKA recovery, they differ in their primary focus and clinical applicability. Gait speed and SCTs assess mobility in dynamic movement tasks, while STS focuses on lower limb muscle strength in a controlled, stationary task. Gait speed tests are most effective for evaluating overall mobility and endurance, whereas SCTs are ideal for assessing functional strength and dynamic balance [83]. In contrast, STS tests specifically target muscle power and endurance without assessing walking ability or balance [84]. Despite these differences, all three tests share a common goal of evaluating lower limb function and have been shown to be reliable, valid, and responsive to post-TKA functional improvements [78]. However, it is worth noting that both gait speed and SCTs may present logistical challenges in smaller clinical settings due to space constraints and increased time requirements for setup and execution, potentially limiting their routine use in busy outpatient environments.

Dobson et al [78] put forth a set of recommended performance-based measures for individuals diagnosed with hip and knee osteoarthritis or following joint replacement. This selection was based on expert surveys and systematic reviews, considering the evidence supporting the measurement properties of the tests, their feasibility, scoring methods, and expert consensus. The resulting recommended measures included the 30-Second Chair Stand Test (30CST), 40 m Fast-paced Walk Test, SCT, Time Up and Go (TUG), and 6-Minute Walk Test (6MWT). In a subsequent update by Dobson et al [85], the 10 Meter Fast-paced Walk Test and 20-second stair climb test were suggested as alternatives due to complexities in administering the 40-meter Fast-paced Walk and scoring the 11 Step Stair Climb Test.

A task force led by Westby et al [86] created the Total Joint Arthroplasty and Outcome Measures toolkit to be used before and after arthroplasty, which included: 30-second STS, gait speed, stair climb test, single leg stance test, 6MWT, TUG, and functional reach. Prior to the creation of this toolkit, Zeni et al [87] developed the Delaware Osteoarthritis Profile, a comprehensive set of tests that have been effectively used to measure functional performance pre and post knee arthroplasty, which include: TUG, SCT, and 6MWT.

While the literature contains numerous tests, the focus here will be on muscle strength performance tests. STS with its versions has been identified by Bergquist et al [88] in their systematic literature review to be the most appropriate performance-based clinical muscle strength test.

The 30CST is a component of the Senior Fitness Test and involves counting the number of sit-to-stand repetitions completed within a 30-second timeframe. Individuals who are more than halfway through a repetition at the 30-second mark are credited with completing the final repetition. This test has shown good to excellent reliability, with ICCs of approximately 0.84 to 0.92 in clinical populations, and its validity is supported by moderate to strong correlations with leg press strength measures [89].

The 5-Repetition Sit-to-Stand (5R-STS) test has gained widespread use as either a component of the short physical performance battery or as a standalone assessment tool in numerous studies. This test involves performing 5 sit-to-stand repetitions from a standard chair (44‐48 cm height), with timing commencing upon a specific command or the initiation of the first movement and ending when the fifth stand-up is accomplished or when the patient sits down following the fifth stand-up [90]. The reliability of the 5R-STS test has been estimated across 10 studies, yielding a coefficient of 0.81, as reported by Bohannon (2011) [84]. The validity of this test as a measure of lower limb functional muscle strength is supported by its strong correlation with knee extension strength, as demonstrated by Lord et al [82]. Furthermore, the correlation between performance on the 5R-STS test and the TUG test, as well as gait speed tests, adds further support for its validity, as highlighted by Schaubert and Bohannon [91].

Although the 5R-STS and 30CST tests involve identical movements, they are not interchangeable. The 5R-STS indicates lower limb speed and power for those with lower physical function post-TKA, while the 30CST measures lower limb endurance for those with higher physical function [92].

While outcome measures are often viewed as purely psychometric, their clinical relevance is established through their role as important drivers of long-term surgical success. For instance, a recent systematic review and meta-analysis by Sumbal et al [93] demonstrated that loss of muscle mass, or sarcopenia, is a significant risk factor for prosthetic loosening, which remains a primary indication for revision surgery. Furthermore, physical performance levels and quadriceps strength have been shown to be strongly associated with patient satisfaction and restoration of functional ability after TKA [94]. More recently, Akatsuka et al [95] reinforced this clinical relevance by reporting significant correlations between quadriceps muscle mass and postoperative satisfaction, suggesting that these objective measures are predictive of how a patient perceives their surgical result.

Discussion

Principal findings

A range of assessment tools is available to evaluate recovery following TKA, each with distinct advantages and limitations. This review highlights the use, strengths, and constraints of these measures and recommends a comprehensive, multidimensional evaluation framework. Such an approach enhances the accuracy and clinical relevance of recovery assessments by triangulating data across patient-reported outcomes, strength measures, functional performance, and muscle mass.

Among muscle mass assessment techniques, DXA remains the gold standard due to its high precision and reliability. However, alternative methods such as BIA offer a noninvasive, lower-cost option suitable for clinical settings, despite reduced accuracy and sensitivity to minor changes in muscle composition.

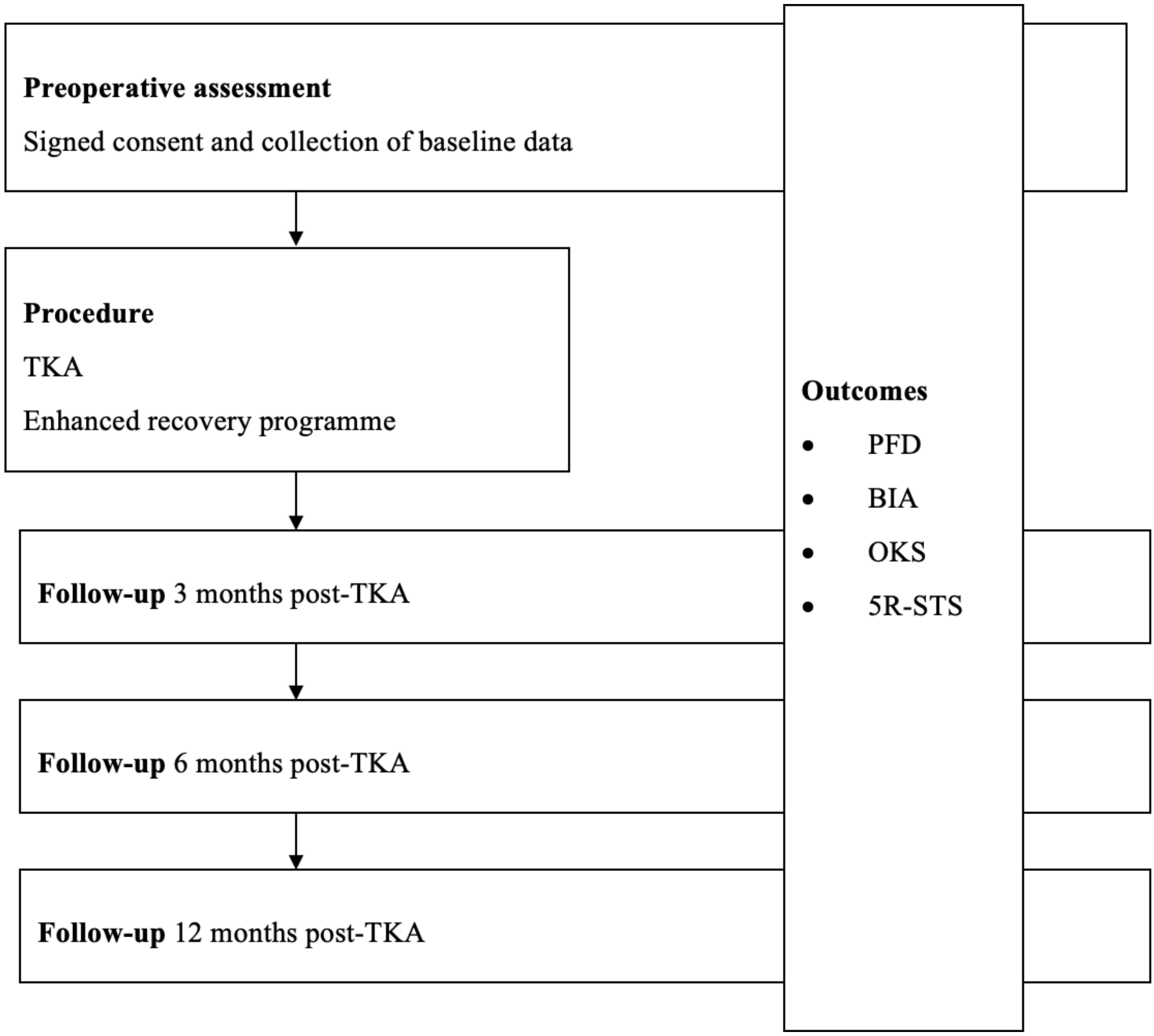
PROMs provide valuable insights into patient perspectives. The OKS is widely employed due to its ease of use, absence of ceiling or floor effects in the short term, and extensive validation in clinical research. KOOS, while broader in scope, often demonstrates ceiling effects and a higher completion burden. WOMAC overlaps conceptually but lacks OKS's responsiveness in early recovery. KSS combines patient and clinician input but is less feasible for routine use due to scoring complexity and reliance on in-person assessment. OKS offers the optimal balance of validity, efficiency, and clinical use in TKA follow-up. Nevertheless, PROMs alone do not provide objective strength assessments, which are critical for a comprehensive evaluation.

In contrast, IKD offers precise quantification of muscle strength, yet its high cost and operational complexity restrict its widespread application. As a practical alternative, PFD provides a reliable, cost-effective, and portable means of assessing muscle strength, making it highly suitable for routine clinical practice.

Performance-based measures such as the 5R-STS test offer objective, time-efficient insights into lower limb function. Compared to the TUG, 6MWT, or SCT, the 5R-STS requires less space, is less influenced by cardiovascular limitations, and is more feasible for patients with early post-op mobility impairments. It also shows strong correlation with knee extension strength and mobility metrics, enhancing its predictive use in post-TKA recovery.

This specific combination, OKS, PFD, BIA, and 5R-STS is what we have recommended in post-TKA follow-up (Figure 2) because it balances psychometric validity, clinical feasibility, and comprehensive recovery profiling. It provides a more complete view of TKA recovery than any individual tool. While alternatives like the 6MWT or SCT yield valuable data, their requirements for time, space, or greater cardiopulmonary reserve make them less practical for routine follow-up or for patients with limited mobility or comorbidities. Likewise, the TUG, though simple, primarily reflects balance and walking ability but lacks sensitivity to muscle power or patient satisfaction domains.

Figure 2. Clinical workflow algorithm for integrating recommended outcome measures throughout the total knee arthroplasty (TKA) pathway. The flowchart outlines the timing and selection of assessment tools from the preoperative assessment appointment to the final follow-up. 5R-STS: 5-Repetition Sit-to-Stand; BIA: bioelectrical impedance analysis; OKS: Oxford Knee Score; PFD: portable fixed dynamometer.



Limitations

Finally, it is important to note that the interpretability and use of these tools may vary across populations. For instance, sarcopenia, obesity, or cardiopulmonary conditions may skew performance-based results or BIA readings, requiring clinicians to contextualize outcomes based on age, BMI, and comorbid burden. By triangulating across these assessments, clinicians can reduce bias from any single measure and develop a more individualized recovery evaluation.

Conclusions

Throughout this paper, evidence suggests that incorporating muscle-based and performance-based measures alongside PROMs is essential for a comprehensive and clinically relevant assessment of TKA recovery. However, future research should ensure there are multicenter trials to validate the integration of the proposed assessments, as well as consensus agreement with clinicians (ie, through a Delphi approach) which translates into guidelines. Additionally, a greater understanding of intrinsic skeletal muscle (mal)adaptations following TKA may provide valuable insight into the mechanisms underlying persistent functional limitations.

Authors' Contributions

Conceptualization: DP, AB
Formal analysis: AB
Investigation: AB
Methodology: AB, DP

Supervision: NM, DP
Visualization: AB
Writing – original draft: AB
Writing – review & writing: AB, VM
All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparative analysis of outcome measures used in total knee arthroplasty (TKA), including patient-reported outcomes, muscle strength, muscle mass, and physical performance tools. Comparison criteria include measurement type, reliability, validity, ease of use, clinical relevance, limitations, and best use case.
[DOCX File, 14 KB - periop_v9i1e84011_app1.docx]

References

1. Neogi T. The epidemiology and impact of pain in osteoarthritis. Osteoarthritis Cartilage 2013 Sep;21(9):1145-1153. [doi: 10.1016/j.joca.2013.03.018] [Medline: 23973124]
2. Vos T, Lim SS, Abbafati C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: Rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006 Aug;20(4):721-740. [doi: 10.1016/j.berh.2006.05.002] [Medline: 16979535]
4. Rissolio L, Sabatini L, Risitano S, et al. Is it the surgeon, the patient, or the device? A comprehensive clinical and radiological evaluation of factors influencing patient satisfaction in 648 total knee arthroplasties. J Clin Med 2021 Jun 12;10(12):2599. [doi: 10.3390/jcm10122599] [Medline: 34204628]
5. Lewis GN, Rice DA, Rashid U, McNair PJ, Kluger MT, Somogyi AA. Trajectories of pain and function outcomes up to 5 to 8 years following total knee arthroplasty. J Arthroplasty 2023 Aug;38(8):1516-1521. [doi: 10.1016/j.arth.2023.02.020] [Medline: 36805116]
6. Tungtrongjit Y, Weingkum P, Saunkool P. The effect of preoperative quadriceps exercise on functional outcome after total knee arthroplasty. J Med Assoc Thai 2012 Oct;95 Suppl 10:S58-S66. [Medline: 23451440]
7. Vajapey SP, Morris J, Spitzer AI, Glassman AH, Greco NJ, Li M. Outcome reporting patterns in total knee arthroplasty: a systematic review. J Clin Orthop Trauma 2020 Jul;11(Suppl 4):S464-S471. [doi: 10.1016/j.jcot.2020.05.014] [Medline: 32774013]
8. Mizner RL, Petterson SC, Clements KE, Zeni JA, Irrgang JJ, Snyder-Mackler L. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. J Arthroplasty 2011 Aug;26(5):728-737. [doi: 10.1016/j.arth.2010.06.004] [Medline: 20851566]
9. Naili JE, Iversen MD, Esbjörnsson AC, et al. Deficits in functional performance and gait one year after total knee arthroplasty despite improved self-reported function. Knee Surg Sports Traumatol Arthrosc 2017 Nov;25(11):3378-3386. [doi: 10.1007/s00167-016-4234-7] [Medline: 27436196]
10. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998 Jan;80(1):63-69. [doi: 10.1302/0301-620x.80b1.7859] [Medline: 9460955]
11. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res 1989 Nov(248):13-14. [Medline: 2805470]
12. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988 Dec;15(12):1833-1840. [Medline: 3068365]
13. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998 Aug;28(2):88-96. [doi: 10.2519/jospt.1998.28.2.88] [Medline: 9699158]
14. Martimbianco ALC, Calabrese FR, Iha LAN, Petrilli M, Lira Neto O, Carneiro Filho M. Reliability of the “American Knee Society Score” (AKSS) [Article in Portuguese]. Acta Ortop Bras 2012;20(1):34-38. [doi: 10.1590/S1413-78522012000100007] [Medline: 24453578]
15. Hodges PW, Smeets RJ. Interaction between pain, movement, and physical activity: short-term benefits, long-term consequences, and targets for treatment. Clin J Pain 2015 Feb;31(2):97-107. [doi: 10.1097/AJP.0000000000000098] [Medline: 24709625]
16. Merkle SL, Sluka KA, Frey-Law LA. The interaction between pain and movement. J Hand Ther 2020;33(1):60-66. [doi: 10.1016/j.jht.2018.05.001] [Medline: 30025839]

17. Ghanem E, Pawasarat I, Lindsay A, et al. Limitations of the Knee Society Score in evaluating outcomes following revision total knee arthroplasty. J Bone Joint Surg Am 2010 Oct 20;92(14):2445-2451. [doi: 10.2106/JBJS.I.00252] [Medline: 20962195]
18. Vogel N, Kaelin R, Rychen T, Wendelspiess S, Müller-Gerbl M, Arnold MP. Comparison of responsiveness of patient-reported outcome measures after total knee arthroplasty. J Arthroplasty 2024 Jun;39(6):1487-1495.e2. [doi: 10.1016/j.arth.2023.11.029] [Medline: 38043745]
19. Bach CM, Nogler M, Steingruber IE, et al. Scoring systems in total knee arthroplasty. Clin Orthop Relat Res 2002 Jun;399(399):184-196. [doi: 10.1097/00003086-200206000-00022] [Medline: 12011708]
20. Karimijashni M, Abtahi F, Abbasalipour S, et al. Functional patient-reported outcome measures after hip or knee arthroplasty: a systematic review and content analysis using the International Classification of Functioning, Disability, and Health. Arthritis Care Res (Hoboken) 2024 Dec;76(12):1703-1722. [doi: 10.1002/acr.25413] [Medline: 39119739]
21. Noble PC, Scuderi GR, Brekke AC, et al. Development of a new Knee Society scoring system. Clin Orthop Relat Res 2012 Jan;470(1):20-32. [doi: 10.1007/s11999-011-2152-z] [Medline: 22065240]
22. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The new Knee Society knee scoring system. Clin Orthop Relat Res 2012 Jan;470(1):3-19. [doi: 10.1007/s11999-011-2135-0] [Medline: 22045067]
23. Siljander MP, McQuivey KS, Fahs AM, Galasso LA, Serdahely KJ, Karadsheh MS. Current trends in patient-reported outcome measures in total joint arthroplasty: a study of 4 major orthopaedic journals. J Arthroplasty 2018 Nov;33(11):3416-3421. [doi: 10.1016/j.arth.2018.06.034] [Medline: 30057269]
24. Ritter MA, Davis KE, Davis P, et al. Preoperative malalignment increases risk of failure after total knee arthroplasty. J Bone Joint Surg Am 2013 Jan 16;95(2):126-131. [doi: 10.2106/JBJS.K.00607] [Medline: 23324959]
25. Capin JJ, Bade MJ, Jennings JM, Snyder-Mackler L, Stevens-Lapsley JE. Total knee arthroplasty assessments should include strength and performance-based functional tests to complement range-of-motion and patient-reported outcome measures. Phys Ther 2022 Jun 3;102(6):pzac033. [doi: 10.1093/ptj/pzac033] [Medline: 35358318]
26. Ben-Shlomo Y, Blom A, Boulton C, et al. The National Joint Registry 19th annual report 2022. : National Joint Registry; 2022 URL: https://www.njrcentre.org.uk/wp-content/uploads/NJR-Supported-Research.pdf [accessed 2025-03-18]
27. Lee MK, Naessens JM, Eton DT, O'Byrne TJ, Nyman MA. Functional outcomes and health-related quality of life before and after primary total knee replacement for patients from diverse geographic regions. J Arthroplasty 2021 Jun;36(6):1958-1965. [doi: 10.1016/j.arth.2021.01.043] [Medline: 33581972]
28. Beard DJ, Knezevic K, Al-Ali S, Dawson J, Price AJ. The use of outcome measures relating to the knee. Orthop Trauma 2010 Aug;24(4):309-316. [doi: 10.1016/j.mporth.2010.04.001]
29. Harris K, Lim CR, Dawson J, Fitzpatrick R, Beard DJ, Price AJ. The Oxford Knee Score and its subscales do not exhibit a ceiling or a floor effect in knee arthroplasty patients: an analysis of the National Health Service PROMs data set. Knee Surg Sports Traumatol Arthrosc 2017 Sep;25(9):2736-2742. [doi: 10.1007/s00167-015-3788-0] [Medline: 26519187]
30. Marx RG, Jones EC, Atwan NC, Closkey RF, Salvati EA, Sculco TP. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. J Bone Joint Surg Am 2005 Sep;87(9):1999-2005. [doi: 10.2106/JBJS.D.02286] [Medline: 16140815]
31. Beard DJ, Davies LJ, Cook JA, et al. The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial. Lancet 2019 Aug 31;394(10200):746-756. [doi: 10.1016/S0140-6736(19)31281-4] [Medline: 31326135]
32. Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012 Feb;20(2):87-92. [doi: 10.1016/j.joca.2011.11.007] [Medline: 22155074]
33. Berth A, Urbach D, Awiszus F. Improvement of voluntary quadriceps muscle activation after total knee arthroplasty. Arch Phys Med Rehabil 2002 Oct;83(10):1432-1436. [doi: 10.1053/apmr.2002.34829] [Medline: 12370881]
34. Bellamy N. WOMAC: A 20-year experiential review of a patient-centered self-reported health status questionnaire. J Rheumatol 2002 Dec;29(12):2473-2476. [Medline: 12465137]
35. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS)—validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003 May 25;1(1):17. [doi: 10.1186/1477-7525-1-17] [Medline: 12801417]
36. Antosh IJ, Svoboda SJ, Peck KY, Garcia EJ, Cameron KL. Change in KOOS and WOMAC scores in a young athletic population with and without anterior cruciate ligament injury. Am J Sports Med 2018 Jun;46(7):1606-1616. [doi: 10.1177/0363546518768753] [Medline: 29733680]
37. Collins NJ, Prinsen CAC, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthritis Cartilage 2016 Aug;24(8):1317-1329. [doi: 10.1016/j.joca.2016.03.010] [Medline: 27012756]
38. Wang Y, Yin M, Zhu S, Chen X, Zhou H, Qian W. Patient-reported outcome measures used in patients undergoing total knee arthroplasty. Bone Joint Res 2021 Mar;10(3):203-217. [doi: 10.1302/2046-3758.103.BJR-2020-0268.R1] [Medline: 33734821]

39. Harris K, Dawson J, Gibbons E, et al. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. Patient Relat Outcome Meas 2016;7:101-108. [doi: 10.2147/PROM.S97774] [Medline: 27524925]
40. Alviar MJ, Olver J, Brand C, et al. Do patient-reported outcome measures in hip and knee arthroplasty rehabilitation have robust measurement attributes? A systematic review. J Rehabil Med 2011 Jun;43(7):572-583. [doi: 10.2340/16501977-0828] [Medline: 21607295]
41. Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a “good” measure. Clin Geriatr Med 2012 Aug;28(3):367-394. [doi: 10.1016/j.cger.2012.05.007] [Medline: 22840304]
42. Gerrits KH, Beltman MJ, Koppe PA, et al. Isometric muscle function of knee extensors and the relation with functional performance in patients with stroke. Arch Phys Med Rehabil 2009 Mar;90(3):480-487. [doi: 10.1016/j.apmr.2008.09.562] [Medline: 19254615]
43. Maffiuletti NA. Assessment of hip and knee muscle function in orthopaedic practice and research. J Bone Joint Surg Am 2010 Jan;92(1):220-229. [doi: 10.2106/JBJS.I.00305] [Medline: 20048117]
44. Hyun CW, Han EY, Im SH, et al. Hemiparetic knee extensor strength and balance function are predictors of ambulatory function in subacute stroke patients. Ann Rehabil Med 2015 Aug;39(4):577-585. [doi: 10.5535/arm.2015.39.4.577] [Medline: 26361594]
45. Norris R, Morrison S, Price A, et al. Inline dynamometry provides reliable measurements of quadriceps strength in healthy and ACL-reconstructed individuals and is a valid substitute for isometric electromechanical dynamometry following ACL reconstruction. Knee 2024 Jan;46:136-147. [doi: 10.1016/j.knee.2023.12.006] [Medline: 38142660]
46. Lienhard K, Lauermann SP, Schneider D, Item-Glatthorn JF, Casartelli NC, Maffiuletti NA. Validity and reliability of isometric, isokinetic and isoinertial modalities for the assessment of quadriceps muscle strength in patients with total knee arthroplasty. J Electromyogr Kinesiol 2013 Dec;23(6):1283-1288. [doi: 10.1016/j.jelekin.2013.09.004] [Medline: 24113423]
47. Ieiri A, Tushima E, Ishida K, Inoue M, Kanno T, Masuda T. Reliability of measurements of hip abduction strength obtained with a hand-held dynamometer. Physiother Theory Pract 2015 Feb;31(2):146-152. [doi: 10.3109/09593985.2014.960539] [Medline: 25264015]
48. Bohannon RW, Kindig J, Sabo G, Duni AE, Cram P. Isometric knee extension force measured using a handheld dynamometer with and without belt-stabilization. Physiother Theory Pract 2012 Oct;28(7):562-568. [doi: 10.3109/09593985.2011.640385] [Medline: 22191509]
49. Bohannon RW, Pritchard RO, Glenney SS. Portable belt–stabilized hand-held dynamometry set-up for measuring knee extension force. Isokinet Exerc Sci 2013 Nov;21(4):325-329. [doi: 10.3233/IES-130517]
50. Martins J, da Silva JR, da Silva MRB, Bevilaqua-Grossi D. Reliability and validity of the belt-stabilized handheld dynamometer in hip- and knee-strength Tests. J Athl Train 2017 Sep;52(9):809-819. [doi: 10.4085/1062-6050-52.6.04] [Medline: 28787180]
51. Toonstra J, Mattacola CG. Test-retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. J Sport Rehabil 2013 Feb 1;22(1):10.1123/jsr.2013.TR7. [doi: 10.1123/jsr.2013.TR7] [Medline: 22951307]
52. Abdalla PP, Carvalho ADS, Dos Santos AP, et al. One-repetition submaximal protocol to measure knee extensor muscle strength among older adults with and without sarcopenia: a validation study. BMC Sports Sci Med Rehabil 2020;12(1):29. [doi: 10.1186/s13102-020-00178-9] [Medline: 32391159]
53. Horstmann H, Medico P, Lasch F, Krutsch W, Weber-Spickschen TS. Simplified measurement of maximum strength after knee surgery: application-based knee-training device compared to isokinetic testing. Open Access J Sports Med 2020;11:43-49. [doi: 10.2147/OAJSM.S214598] [Medline: 32158284]
54. Whiteley R, Jacobsen P, Prior S, Skazalski C, Otten R, Johnson A. Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. J Sci Med Sport 2012 Sep;15(5):444-450. [doi: 10.1016/j.jsams.2012.01.003] [Medline: 22424705]
55. Lauermann SP, Lienhard K, Item-Glatthorn JF, Casartelli NC, Maffiuletti NA. Assessment of quadriceps muscle weakness in patients after total knee arthroplasty and total hip arthroplasty: methodological issues. J Electromyogr Kinesiol 2014 Apr;24(2):285-291. [doi: 10.1016/j.jelekin.2013.10.018] [Medline: 24290027]
56. Tsukasaki K, Matsui Y, Arai H, et al. Association of muscle strength and gait speed with cross-sectional muscle area determined by mid-thigh computed tomography—a comparison with skeletal muscle mass measured by dual-energy X-ray absorptiometry. J Frailty Aging 2020;9(2):82-89. [doi: 10.14283/jfa.2020.16] [Medline: 32259181]
57. Riviati N, Indra B. Relationship between muscle mass and muscle strength with physical performance in older adults: a systematic review. SAGE Open Med 2023;11:20503121231214650. [doi: 10.1177/20503121231214650] [Medline: 38033420]
58. Heymsfield SB, Adamek M, Gonzalez MC, Jia G, Thomas DM. Assessing skeletal muscle mass: historical overview and state of the art. J Cachexia Sarcopenia Muscle 2014 Mar;5(1):9-18. [doi: 10.1007/s13539-014-0130-5] [Medline: 24532493]
59. Binay Safer V, Safer U. Usefulness and limitations of single-slice computed tomography analysis at the third lumbar region in the assessment of sarcopenia. Crit Care 2013 Nov 20;17(6):466. [doi: 10.1186/cc13123] [Medline: 24252497]

60. Gibson DJ, Burden ST, Strauss BJ, Todd C, Lal S. The role of computed tomography in evaluating body composition and the influence of reduced muscle mass on clinical outcome in abdominal malignancy: a systematic review. Eur J Clin Nutr 2015 Oct;69(10):1079-1086. [doi: 10.1038/ejcn.2015.32] [Medline: 25782424]
61. Lemos T, Gallagher D. Current body composition measurement techniques. Curr Opin Endocrinol Diabetes Obes 2017 Oct;24(5):310-314. [doi: 10.1097/MED.0000000000000360] [Medline: 28696961]
62. Strandberg S, Wretling ML, Wredmark T, Shalabi A. Reliability of computed tomography measurements in assessment of thigh muscle cross-sectional area and attenuation. BMC Med Imaging 2010 Aug 11;10(1):18. [doi: 10.1186/1471-2342-10-18] [Medline: 20701775]
63. Erlandson MC, Lorbergs AL, Mathur S, Cheung AM. Muscle analysis using pQCT, DXA and MRI. Eur J Radiol 2016 Aug;85(8):1505-1511. [doi: 10.1016/j.ejrad.2016.03.001] [Medline: 27005009]
64. Heymsfield SB, Gonzalez MC, Lu J, Jia G, Zheng J. Skeletal muscle mass and quality: evolution of modern measurement concepts in the context of sarcopenia. Proc Nutr Soc 2015 Nov;74(4):355-366. [doi: 10.1017/S0029665115000129] [Medline: 25851205]
65. Prado CMM, Heymsfield SB. Lean tissue imaging: a new era for nutritional assessment and intervention. JPEN J Parenter Enteral Nutr 2014 Nov;38(8):940-953. [doi: 10.1177/0148607114550189] [Medline: 25239112]
66. Lustgarten MS, Fielding RA. Assessment of analytical methods used to measure changes in body composition in the elderly and recommendations for their use in phase II clinical trials. J Nutr Health Aging 2011 May;15(5):368-375. [doi: 10.1007/s12603-011-0049-x] [Medline: 21528163]
67. Maden-Wilkinson TM, Degens H, Jones DA, McPhee JS. Comparison of MRI and DXA to measure muscle size and age-related atrophy in thigh muscles. J Musculoskelet Neuronal Interact 2013 Sep;13(3):320-328. [Medline: 23989253]
68. Kim J, Wang Z, Heymsfield SB, Baumgartner RN, Gallagher D. Total-body skeletal muscle mass: estimation by a new dual-energy X-ray absorptiometry method. Am J Clin Nutr 2002 Aug;76(2):378-383. [doi: 10.1093/ajcn/76.2.378] [Medline: 12145010]
69. Taylor AE, Kuper H, Varma RD, et al. Validation of dual energy X-ray absorptiometry measures of abdominal fat by comparison with magnetic resonance imaging in an Indian population. PLoS One 2012;7(12):e51042. [doi: 10.1371/journal.pone.0051042] [Medline: 23272086]
70. Hangartner TN, Warner S, Braillon P, Jankowski L, Shepherd J. The official positions of the International Society for Clinical Densitometry: acquisition of dual-energy X-ray absorptiometry body composition and considerations regarding analysis and repeatability of measures. J Clin Densitom 2013;16(4):520-536. [doi: 10.1016/j.jocd.2013.08.007] [Medline: 24183641]
71. Toombs RJ, Ducher G, Shepherd JA, De Souza MJ. The impact of recent technological advances on the trueness and precision of DXA to assess body composition. Obesity (Silver Spring) 2012 Jan;20(1):30-39. [doi: 10.1038/oby.2011.211] [Medline: 21760631]
72. Johnson Stoklossa CA, Forhan M, Padwal RS, Gonzalez MC, Prado CM. Practical considerations for body composition assessment of adults with class ii/iii obesity using bioelectrical impedance analysis or dual-energy x-ray absorptiometry. Curr Obes Rep 2016 Dec;5(4):389-396. [doi: 10.1007/s13679-016-0228-5] [Medline: 27591783]
73. Khalil SF, Mohktar MS, Ibrahim F. The theory and fundamentals of bioimpedance analysis in clinical status monitoring and diagnosis of diseases. Sensors (Basel) 2014 Jun 19;14(6):10895-10928. [doi: 10.3390/s140610895] [Medline: 24949644]
74. Marra M, Sammarco R, De Lorenzo A, et al. Assessment of body composition in health and disease using bioelectrical impedance analysis (BIA) and dual energy x-ray absorptiometry (DXA): a critical overview. Contrast Media Mol Imaging 2019;2019:3548284. [doi: 10.1155/2019/3548284] [Medline: 31275083]
75. Buckinx F, Reginster JY, Dardenne N, et al. Concordance between muscle mass assessed by bioelectrical impedance analysis and by dual energy x-ray absorptiometry: a cross-sectional study. BMC Musculoskelet Disord 2015 Mar 18;16(1):60. [doi: 10.1186/s12891-015-0510-9] [Medline: 25887598]
76. Ling CHY, de Craen AJM, Slagboom PE, et al. Accuracy of direct segmental multi-frequency bioimpedance analysis in the assessment of total body and segmental body composition in middle-aged adult population. Clin Nutr 2011 Oct;30(5):610-615. [doi: 10.1016/j.clnu.2011.04.001] [Medline: 21555168]
77. Jette DU, Hunter SJ, Burkett L, et al. Physical therapist management of total knee arthroplasty. Phys Ther 2020 Aug 31;100(9):1603-1631. [doi: 10.1093/ptj/pzaa099] [Medline: 32542403]
78. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013 Aug;21(8):1042-1052. [doi: 10.1016/j.joca.2013.05.002] [Medline: 23680877]
79. Rehabilitative care best practices for patients with primary hip & knee replacements. : Rehabilitative Care Alliance; 2017 URL: https://rehabcarealliance.ca/wp-content/uploads/2025/03/TJR_Framework.pdf [accessed 2026-02-13]
80. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA 2011 Jan 5;305(1):50-58. [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
81. Fenner V, Behrend H, Kuster MS. Whole body gait function during stair ascending and level walking in patients following total knee arthroplasty. Int J Phys Med Rehabil 2014;02(5). [doi: 10.4172/2329-9096.1000231]

82. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci 2002 Aug;57(8):M539-M543. [doi: 10.1093/gerona/57.8.m539] [Medline: 12145369]
83. Mizner RL, Petterson SC, Snyder-Mackler L. Quadriceps strength and the time course of functional recovery after total knee arthroplasty. J Orthop Sports Phys Ther 2005 Jul;35(7):424-436. [doi: 10.2519/jospt.2005.35.7.424] [Medline: 16108583]
84. Bohannon RW. Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults. J Strength Cond Res 2011 Nov;25(11):3205-3207. [doi: 10.1519/JSC.0b013e318234e59f] [Medline: 21904240]
85. Dobson F, Hinman RS, Hall M, et al. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. Osteoarthritis Cartilage 2017 Nov;25(11):1792-1796. [doi: 10.1016/j.joca.2017.06.006] [Medline: 28647467]
86. Hoens A, Westby M, Longstaff S, Duggan M. Total Joint Arthroplasty and Outcome Measures (TJAOM) toolkit. The University of British Columbia. 2016. URL: https://physicaltherapy.med.ubc.ca/physical-therapy-knowledge-broker/tjaom-toolkit/ [accessed 2026-02-13]
87. Zeni JA, Axe MJ, Snyder-Mackler L. Clinical predictors of elective total joint replacement in persons with end-stage knee osteoarthritis. BMC Musculoskelet Disord 2010 May 6;11(1):86. [doi: 10.1186/1471-2474-11-86] [Medline: 20459622]
88. Bergquist R, Weber M, Schwenk M, et al. Performance-based clinical tests of balance and muscle strength used in young seniors: a systematic literature review. BMC Geriatr 2019 Jan 9;19(1):9. [doi: 10.1186/s12877-018-1011-0] [Medline: 30626340]
89. Gill S, Hely R, Page RS, Hely A, Harrison B, Landers S. Thirty second chair stand test: test-retest reliability, agreement and minimum detectable change in people with early-stage knee osteoarthritis. Physiother Res Int 2022 Jul;27(3):e1957. [doi: 10.1002/pri.1957] [Medline: 35592902]
90. Klukowska AM, Staartjes VE, Vandertop WP, Schröder ML. Five-repetition sit-to-stand test performance in healthy individuals: reference values and predictors from 2 prospective cohorts. Neurospine 2021 Dec;18(4):760-769. [doi: 10.14245/ns.2142750.375] [Medline: 35000330]
91. Schaubert KL, Bohannon RW. Reliability and validity of three strength measures obtained from community-dwelling elderly persons. J Strength Cond Res 2005 Aug;19(3):717-720. [doi: 10.1519/R-15954.1] [Medline: 16095431]
92. McCarthy EK, Horvat MA, Holtsberg PA, Wisenbaker JM. Repeated chair stands as a measure of lower limb strength in sexagenarian women. J Gerontol A Biol Sci Med Sci 2004 Nov;59(11):1207-1212. [doi: 10.1093/gerona/59.11.1207] [Medline: 15602077]
93. Sumbal R, Abbas M, Sheikh SM, Sumbal A. Prevalence and clinical impact of sarcopenia in patients undergoing total joint arthroplasty: a systematic review and a meta-analysis. J Arthroplasty 2024 Dec;39(12):3128-3135. [doi: 10.1016/j.arth.2024.06.021] [Medline: 38901711]
94. Furu M, Ito H, Nishikawa T, et al. Quadriceps strength affects patient satisfaction after total knee arthroplasty. J Orthop Sci 2016 Jan;21(1):38-43. [doi: 10.1016/j.jos.2015.10.002] [Medline: 26755384]
95. Akatsuka Y, Teramoto A, Takashima H, Okada Y, Watanabe K, Yamashita T. Relationships of cross-sectional area of the thigh muscles before or after total knee arthroplasty with postoperative pain or patient satisfaction: a retrospective, exploratory study. Asia Pac J Sports Med Arthrosc Rehabil Technol 2023 Jul;33:20-24. [doi: 10.1016/j.asmart.2023.08.006] [Medline: 37680195]

Abbreviations

1-RM: 1–repetition maximum
30CST: 30-Second Chair Stand Test
5R-STS: 5-Repetition Sit-to-Stand
6MWT: 6-Minute Walk Test
ADL: activities of daily living
BIA: bioelectrical impedance analysis
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
CT: computed tomography
DXA: dual-energy X-ray absorptiometry
ICC: intraclass correlation coefficient
KOOS: Knee Injury and Osteoarthritis Outcome Score
KSS: Knee Society Clinical Rating System
MRI: magnetic resonance imaging
OKS: Oxford Knee Score
PFD: portable fixed dynamometer
PROM: patient-related outcome measure
ROM: range of motion

SCT: Stair Climb Test
STS: Sit-to-Stand
TKA: total knee arthroplasty
TUG: Time Up and Go
WOMAC: Western Ontario and McMaster Universities Osteoarthritis

Edited by N Rohatgi; submitted 12.Sep.2025; peer-reviewed by J Most; revised version received 30.Jan.2026; accepted 04.Feb.2026; published 25.Feb.2026.

Please cite as:
Boukabache A, Maruthainar N, Manhas V, Player D
Multidimensional Assessment of Recovery After Total Knee Arthroplasty in Clinical Practice: Critical Narrative Review
JMIR Perioper Med 2026;9:e84011
URL: https://periop.jmir.org/2026/1/e84011
doi:10.2196/84011

Forced-Air Warming Temperature Settings for Treating Postoperative Hypothermia in the Postanesthesia Care Unit: Randomized Controlled Trial

Koravee Pasutharnchat, MD; Rattaphol Seangrung, MD; Sirikarn Sirisophaphong, MD; Wilailuck Wongkum, BNS

Department of Anesthesiology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, 270 Rama 6 Road, Ratchathewi, Bangkok, Thailand

Corresponding Author:
Koravee Pasutharnchat, MD
Department of Anesthesiology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, 270 Rama 6 Road, Ratchathewi, Bangkok, Thailand

Abstract

Background: Hypothermia, defined as a core body temperature below 36 °C, is a common postoperative complication associated with adverse outcomes, including delayed wound healing, infections, and increased bleeding.

Objective: This randomized controlled trial evaluated the efficacy of different forced-air warming system temperature settings in treating postoperative hypothermia in the postanesthesia care unit.

Methods: A total of 132 patients undergoing elective surgery at Ramathibodi Hospital between April 2023 and May 2024 were randomized into 3 groups (n=44 per group): group C (warming set to 38 °C), group F1 (warming set to 42 °C), and group F2 (warming set to 42 °C, reduced to 38 °C after achieving 36 °C). Tympanic temperature was recorded at 5-minute intervals during rewarming and every 10 minutes after normothermia (≥36 °C) was achieved. The primary outcome was rewarming time. Secondary outcomes included the incidence of temperature drops, hemodynamic parameters, adverse events, and patient comfort scores.

Results: Baseline characteristics and clinical variables, including vital signs, were comparable among groups (*P*>.05). Group F2 achieved the shortest mean rewarming time of 33.3 (SD 13.81) min; however, differences between groups were not statistically significant (*P*=.460). Group F2 had the lowest incidence of temperature drops below 36 °C after normothermia (1/44, 2.27%; *P*=.009). Group C had the highest incidence of rewarming exceeding 1 hour (10/44, 22.73%; *P*=.017).

Conclusions: While rewarming times were similar across groups, the protocol using an initial setting of 42 °C followed by a reduction to 38 °C (group F2) effectively minimized temperature drops after normothermia, suggesting its superiority for managing postoperative hypothermia in the postanesthesia care unit.

Trial Registration: Thaiclinicaltrials.org TCTR20231012004; https://www.thaiclinicaltrials.org/show/TCTR20231012004

(JMIR Perioper Med 2026;9:e85045) doi:10.2196/85045

KEYWORDS

postoperative hypothermia; forced-air warming; effectiveness; postanesthesia care unit; temperature settings

Introduction

Hypothermia, defined as a core body temperature below 36 °C, is a frequent complication in patients undergoing elective surgery [1]. Intraoperative hypothermia, if uncorrected, often leads to postoperative hypothermia, particularly in the recovery room, where insufficient warming measures can exacerbate the condition. The prevalence of postoperative hypothermia in the postanesthesia care unit (PACU) has been reported to range from 20% to 28% at arrival and from 18.5% to 26% within 30 minutes after arrival [2]. Postoperative hypothermia is clinically significant, as it has been associated with impaired wound healing, increased risk of surgical site infection, greater blood loss, cardiac arrhythmias, and prolonged hospitalization [3,4]. These adverse consequences highlight the importance of effective temperature management strategies throughout the perioperative period. Recent guidelines and reviews, including Enhanced Recovery After Surgery pathways and the clinical recommendations from the Royal College of Anesthesiologists of Thailand, emphasize the critical role of maintaining normothermia as a core component to reduce surgical site infections and hospital stay [5-8].

Active warming techniques, particularly forced-air warming (FAW), are widely implemented to reduce the incidence of perioperative hypothermia. FAW devices deliver warmed air (32 °C - 47 °C) through a specialized blanket, with built-in safety mechanisms to prevent overheating [9,10]. Systematic reviews have demonstrated that FAW is superior to conventional blankets, reducing the time to restore normothermia by more than an hour [11]. While these findings confirm its effectiveness in facilitating rewarming, the literature remains inconclusive regarding the optimal temperature setting for postoperative use. Most previous studies have focused on preoperative or intraoperative warming [12-18], whereas evidence for postoperative FAW application remains limited. Xu et al [19] reported that FAW at 42 °C was more effective than at 38 °C or conventional blankets in elderly patients undergoing joint replacement. However, the generalizability of that study was restricted by the narrow patient population, short operative times, and limited assessment of adverse events.

At our institution, the prevalence of postoperative hypothermia has remained notable despite the routine availability of FAW systems. Pisitsak et al [20] documented hypothermia in 20% of patients under regional anesthesia and in 16% under general anesthesia. More recent institutional data from 2019 to 2022 indicate an incidence of 23% among surgical patients recovering in the PACU. Furthermore, between 2022 and 2024, the prevalence of hypothermia ranged from 10.8% to 13.8% despite widespread FAW use across multiple surgical specialties, including general surgery, orthopedics, otolaryngology, obstetrics and gynecology, and cardiac surgery (Department of Anesthesiology, Faculty of Medicine Ramathibodi Hospital, Mahidol University. Internal statistical data analyzed via Power BI dashboard, unpublished data, January 2025). These findings suggest that, in addition to patient- and procedure-related factors, variability in FAW temperature settings contributes to inconsistent outcomes.

Current practice in our PACUs uses FAW with adjustable temperature settings ranging from 38 °C to 42 °C; however, no standardized protocol exists to guide optimal temperature selection. This variability reflects broader uncertainty regarding the most effective strategy for postoperative rewarming and underscores the need for evidence-based guidance. To our knowledge, no prior randomized trial has evaluated a step-down temperature protocol (42 °C to 38 °C) in a mixed adult surgical population. By addressing this gap, the present study examines the effectiveness of different FAW temperature settings to inform a pragmatic and standardized PACU warming approach, with the goal of improving consistency in clinical practice and enhancing patient safety.

Methods

Study Design

This study was designed as a prospective randomized controlled trial.

Patients

A total of 132 patients scheduled for elective surgery across various specialties, including general surgery, orthopedics, urology, otolaryngology, obstetrics and gynecology, and cardiac surgery, were enrolled between April 2023 and May 2024. The inclusion criteria consisted of patients aged 18 to 80 years, American Society of Anesthesiologists (ASA) physical status I to III, who were undergoing elective procedures under either general or regional anesthesia, with an expected operating time of at least 2 hours.

Exclusion criteria included patients with a core temperature exceeding 37.5 °C, evidence of infection (eg, sepsis), conditions precluding the use of forced-air warming (eg, burns, agitation, or delirium), those unable to communicate or complete the trial questionnaire, and patients who declined participation.

Sample Size Calculation

A priori sample size calculation was conducted to ensure adequate statistical power for the study's primary outcome: the duration of forced-air warming required for a patient's core temperature to reach ≥36 °C. Based on a previous randomized controlled trial by Xu et al [19], utilizing a 2-sided significance level (α=.05), adjusted for multiple comparisons among the 3 groups (α/3=.017), corresponding to a *z* score of 2.41, a statistical power of 80% (*z* for β=0.84), an estimated SD of 6.45 minutes, and a clinically meaningful difference in rewarming time of 5 minutes, the calculation determined that 36 participants were required per group. To accommodate an anticipated 20% participant dropout rate, the sample size was prudently inflated to 44 participants for each of the 3 intervention groups. This resulted in a total sample size of 132 participants, ensuring robust statistical inference for our findings.

Randomization

Randomization was performed using stratified block randomization with proportional allocation based on the type of anesthesia (general vs regional) to ensure balanced distribution of thermoregulatory impairment mechanisms across groups. A research assistant not involved in patient recruitment generated the computer-based random sequence using permuted blocks of variable size. Allocation was concealed using sequentially numbered, sealed opaque envelopes, which were opened only after participant enrollment. The study personnel responsible for enrollment were different from those assigning participants to groups to ensure the integrity of allocation concealment.

Rewarming

Intraoperative management followed our institution's standard of care, which included the routine use of fluid warmers, application of forced-air warming blankets, and continuous core temperature monitoring for all patients. Upon arrival at the PACUs, patients who met the preliminary criteria were assessed. Only those with a core temperature lower than 36 °C were enrolled and randomly allocated into 3 groups (n=44 per group): group C (forced-air warming set to 38 °C), group F1 (forced-air warming set to 42 °C), and group F2 (forced-air warming initially set to 42 °C, then reduced to 38 °C once the core temperature reached 36 °C).

All participants received identical warming systems and core temperature monitoring devices at the PACU. Rewarming was carried out using a forced-air warming system (Bair Hugger) with a blanket and a core temperature measurement device (Braun ThermoScan ear thermometer).

The rewarming process was monitored and recorded every 5 minutes during the active warming phase. In groups C and F1, the forced-air warmer was discontinued once the core temperature reached ≥36 °C, at which point patients were covered with a regular blanket and monitored every 10 minutes. In group F2, the setting was reduced to 38 °C upon reaching a core temperature of ≥36 °C, and patients were similarly monitored every 10 minutes until discharge from the PACU. Rewarming time was calculated as the time taken for the core temperature to rise from baseline to ≥36 °C, measured in minutes.

Outcome Measures

The primary outcome was the rewarming time, defined as the duration from the initiation of rewarming to the recovery of normothermia (core temperature ≥36 °C). Additionally, the incidence of a decrease in core temperature after achieving normothermia was recorded in each group.

Secondary outcomes included the incidence of adverse events—such as hypotension, hypertension, arrhythmias, nausea or vomiting, pain, and shivering—and patient satisfaction. Patient satisfaction was evaluated using 2 validated instruments: the 5-point Patient Comfort Scale, which measures overall comfort and satisfaction, and the 7-point Thermal Comfort Scale, which assesses subjective thermal sensation ranging from –3 (cold) to +3 (hot), with 0 representing thermal neutrality.

Data Collection

Preoperative and intraoperative data were collected, including patient demographics, surgical procedure, operative time, anesthetic technique, blood loss, and fluid and blood product administration. Upon PACU admission, core temperature was recorded every 5 minutes during the rewarming phase by trained PACU nurses. To ensure consistency, the same nurse performed all assessments for a given patient using the same device and the same ipsilateral ear. Blinding of these nurses was not feasible because the FAW device displayed temperature settings during operation; consequently, the nurses were aware of group allocation, although patients remained blinded. Once the core temperature reached ≥36 °C, measurements continued every 10 minutes until discharge based on the Modified Aldrete scoring system. Throughout the PACU stay, adverse events were monitored continuously, and patient comfort and thermal comfort scores were assessed by the nurses at the time of discharge.

Statistical Analysis

Data were analyzed using SPSS software version 27 (IBM Corp.). Continuous variables were expressed as mean (SD) or median (IQR), depending on the distribution, which was assessed using the Shapiro-Wilk test. Categorical variables were presented as counts and percentages.

For comparisons between groups, one-way ANOVA was used for normally distributed continuous variables, while the Kruskal-Wallis test was applied to non-normally distributed data. Post hoc analyses were performed using Tukey honest significant difference test for ANOVA and Dunn test for the Kruskal-Wallis test, as appropriate. Categorical variables were compared using the chi-square test or Fisher exact test, as required.

Monte Carlo simulation was utilized for the Fisher exact test extension in contingency tables larger than 2×2 where cell counts were sparse (expected count <5), ensuring robust *P* value estimation without violating asymptotic assumptions. Relative risks with 95% CI were reported for significant categorical outcomes. For continuous variables, effect sizes were expressed as Cohen *d* to ensure consistency and enhance clinical interpretability. This was an intention-to-treat analysis, and all randomized patients were analyzed in their assigned groups. A *P* value of <.05 was considered statistically significant.

Ethical Considerations

This study was approved by the Human Research Ethics Committees of Ramathibodi Hospital, Mahidol University (approval number MURA2023/202) and registered at Thaiclinicaltrials.org on March 14, 2023 (approval number TCTR20231012004). Written informed consent was obtained from all participants before enrollment. Participants' privacy and confidentiality were strictly protected, and all data were deidentified before analysis. No financial compensation was provided to participants for their participation in the study.

Results

Baseline Characteristics

A total of 165 patients were assessed for eligibility between April 2023 and May 2024. Thirty-three patients were excluded (2 with ASA physical status >III and 31 normothermic at PACU admission). Finally, 132 patients were included and equally divided into 3 groups (Figure 1): group C (rewarming set to 38 °C, n=44), group F1 (rewarming set to 42 °C, n=44), and group F2 (rewarming set to 42 °C until reaching a core temperature of 36 °C, then reduced to 38 °C, n=44). The baseline characteristics, including age, gender, ASA physical status, BMI, and underlying diseases, were comparable across all groups, as detailed in Table 1.

Figure 1. Flow diagram of patient enrollment. This diagram shows the screening and allocation process for all patients. ASA: American Society of Anesthesiologists; PACU: postanesthesia care unit.

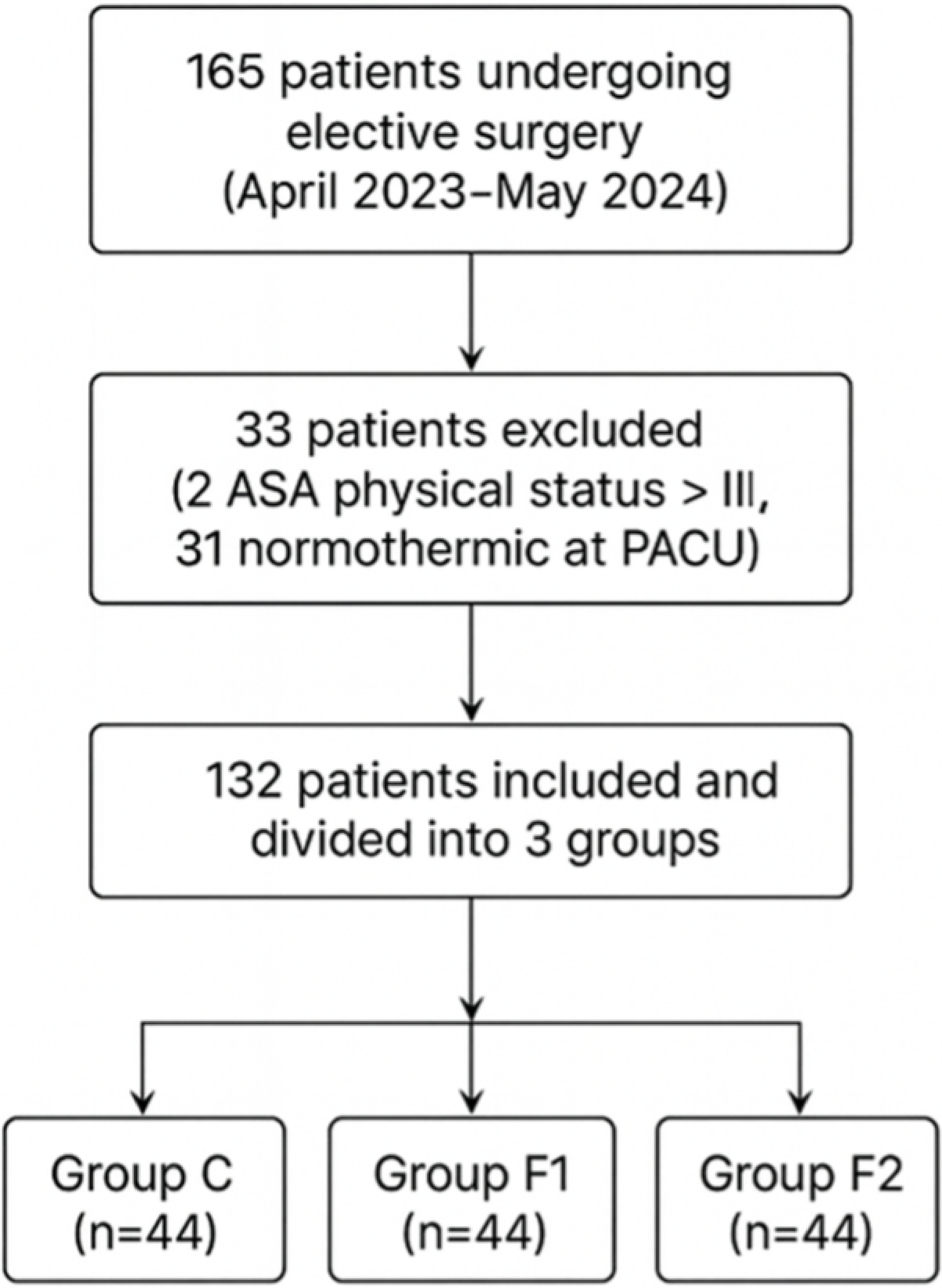


Table . Baseline characteristics of patients in the 3 groups.

Characteristic	Group C (n=44)	Group F1 (n=44)	Group F2 (n=44)	*P* value
Age, n (%)				.265
Elderly (age ≥65 y)	10 (22.73)	13 (29.55)	17 (38.64)	
Nonelderly (age <65 y)	34 (77.27)	31 (70.45)	27 (61.36)	
Gender (male/female)				.083
Male	17	18	9	
Female	27	26	35	
ASA[a] Physical status, n (%)				.965
>2	13 (29.55)	13 (29.55)	14 (31.82)	
≤2	31 (70.45)	31 (70.45)	30 (68.18)	
BMI (kg/m^2), median (IQR)	23.85 (22 - 26.95)	24 (22.43 - 26.3)	23.67 (21.05 - 27)	.738
Underlying diseases, n (%)				
Diabetes mellitus	3 (6.82)	8 (18.18)	7 (15.91)	.259
Hypertension	14 (31.82)	18 (40.91)	19 (43.18)	.511
Obesity	5 (11.36)	6 (13.64)	4 (9.08)	.798
Extreme age	10 (22.73)	12 (27.27)	15 (34.09)	.490
Heart disease	3 (6.82)	4 (9.09)	5 (11.36)	.927
Cerebrovascular disease	3 (6.82)	5 (11.36)	1 (2.27)	.295
Chronic kidney disease	4 (9.09)	5 (11.36)	2 (4.55)	.622
Cancer	5 (11.36)	8 (18.18)	11 (25)	.253
Respiratory disease	2 (4.55)	3 (6.82)	2 (4.55)	>.999
Others	11 (25)	14 (31.82)	7 (15.91)	.217

[a]ASA: American Society of Anesthesiologists.

Intraoperative Data

No significant differences were observed among the 3 groups regarding operative time, type of operation, anesthetic technique, estimated blood loss, total fluid administration, total blood components used, or recorded temperatures, as summarized in Table 2.

Table . Intraoperative parameters of patients in the 3 groups.

Parameter	Group C (n=44)	Group F1 (n=44)	Group F2 (n=44)	*P* value
Operative time (min), median (IQR)	165 (120 - 221.25)	180 (123.75 - 223.75)	195 (137.5 - 233.75)	.153
Anesthetic technique, n (%)				.926
Regional	15 (34.09)	14 (31.82)	14 (31.82)	
General	22 (50)	21 (47.73)	24 (54.55)	
Combined	7 (15.91)	9 (20.45)	6 (13.64)	
Operation, n (%)				.731
Open orthopedic surgery	18 (40.91)	18 (40.91)	17 (38.64)	
Arthroscopic/laparoscopic orthopedic surgery	8 (18.18)	7 (15.91)	11 (25)	
Open gynecologic surgery	3 (6.82)	4 (9.09)	4 (9.09)	
Laparoscopic gynecologic surgery	6 (13.64)	3 (6.82)	8 (18.18)	
Breast surgery	3 (6.82)	0 (0)	2 (4.55)	
Open general surgery	2 (4.55)	2 (4.55)	1 (2.27)	
Urological surgery	1 (2.27)	1 (2.27)	0 (0)	
Thoracic surgery	1 (2.27)	1 (2.27)	0 (0)	
Plastic surgery	18 (40.91)	1 (2.27)	0 (0)	
Vascular surgery	0 (0)	2 (4.55)	0 (0)	
Laparoscopic general surgery	2 (4.55)	4 (9.09)	1 (2.27)	
Otolaryngologic surgery	0 (0)	1 (2.27)	0 (0)	
Intraoperative period				
Large estimated blood loss (≥500 ml), n (%)	1 (2.27)	3 (6.82)	1 (2.27)	.435
Total fluid administered (ml), median (IQR)	975 (762.5 - 1425)	900 (612.5 - 1387.5)	1000 (712.5 - 1487.5)	.808
Total blood component (ml), median (IQR)	0 (0)	0 (0)	0 (0)	.173
Temperature recorded (°C), mean (SD)	36.07 (0.43)	35.93 (0.37)	36.1 (0.32)	.139

Postoperative Data

Postoperative outcomes recorded in the PACU are summarized in Table 3. No significant differences were observed among the 3 groups regarding tympanic temperature upon arrival or hemodynamic parameters, including blood pressure, heart rate, respiratory rate, and SpO_2 (*P*>.05). However, a statistically significant difference was observed in the duration of PACU stay (*P*=.015). Post hoc comparisons revealed significant differences between group C versus group F2 and group F1 versus group F2, indicating a more favorable distribution of discharge times in group F2. Regarding electrocardiogram findings, adverse events were rare, with only 1 patient exhibiting bradycardia.

Table . Postoperative outcomes of patients in the 3 groups.

Parameter	Group C (n=44)	Group F1 (n=44)	Group F2 (n=44)	*P* value
Duration in PACU[a] (min),[c] median (IQR)	60 (60 - 65)	60 (60 - 60)	60 (60 - 60)	.015
Tympanic temperature <36 °C upon arrival in the PACU, n (%)				.165
<35 °C	0 (0)	1 (2.27)	3 (6.82)	
≥35 °C	44 (100)	43 (97.73)	41 (93.18)	
Systolic blood pressure (mmHg), mean (SD)	134.07 (22.57)	134.53 (18.21)	135.95 (21.86)	.910
Diastolic blood pressure (mmHg), mean (SD)	76.32 (31.85)	79.09 (12.5)	94.45 (106.34)	.341
Respiratory rate (per min), mean (SD)	17.07 (3.25)	18.09 (3.2)	17.95 (3.43)	.290
Heart rate (per min), mean (SD)	68.61 (12.12)	72.55 (13.15)	74.75 (13.31)	.081
SpO_2[b] (%), median (IQR)	100 (99 - 100)	100 (99 - 100)	100 (99 - 100)	.688

[a]PACU: postanesthesia care unit.

[b]SpO_2: peripheral capillary oxygen saturation.

[c]A statistically significant difference was observed only in the duration of stay in PACU (*P*=.015; post hoc comparisons revealed significant differences for group C vs group F2 [Cohen *d*=6.29] and group F1 vs group F2 [Cohen *d*=8.55]).

Effect of Different Rewarming Methods

Rewarming outcomes across all dimensions are summarized in Table 4. The time to achieve normothermia is illustrated in Figure 2. Figure 3 shows the number of patients who experienced a drop in core temperature below 36 °C after achieving normothermia, and those who required rewarming for more than 1 hour.

No significant differences were observed among the 3 groups in hemodynamic parameters recorded in the PACU, including systolic and diastolic blood pressure, respiratory rate, heart rate, peripheral capillary oxygen saturation, and electrocardiogram findings (*P*>.05).

While the rewarming time did not differ significantly among the groups, the incidence of patients experiencing a drop in core temperature below 36 °C after achieving normothermia was significantly lower in group F2 compared to groups C and F1 (*P*=.009, Table 4). Patients exhibiting temperature decline required extended thermal support to restore or maintain normothermia. Consequently, a significantly higher proportion of patients in groups C and F1 required active warming for more than 1 hour compared to group F2 (*P*=.017). However, this prolonged warming requirement did not lead to a clinically relevant delay in discharge, as the median duration of PACU stay remained 60 minutes across all groups (Table 3).

Table . Rewarming outcomes of patients in the 3 groups.

Parameter	Group C (n=44)	Group F1 (n=44)	Group F2 (n=44)	*P* value
Mean rewarming time[a] (min), mean (SD)	37.39 (16.58)	35.11 (15.64)	33.30 (13.81)	.460
Decrease temperature below 36 °C after achieving normothermia, n (%)	7 (15.91)	11 (25)	1 (2.27)	.009
Warming more than 1 h, n (%)	10 (22.73)	7 (15.91)	1 (2.27)	.017

[a]For the mean rewarming time, no significant differences were observed among groups; the mean differences (95% CI) compared to group C were −2.28 (−9.03 to 4.47) for group F1 and −4.09 (−10.46 to 2.28) for group F2. However, a statistically significant difference was observed in the proportion of patients with temperature decrease below 36 °C (*P*=.009; significant pairwise differences were observed for group F1 vs group F2 [relative risk (RR)=11.00; 95% CI 1.48 - 81.61]) and those requiring warming for more than 1 h (*P*=.017; significant pairwise differences were observed for group C vs group F2 [RR=10.00; 95% CI 1.34 - 74.84]).

Figure 2. Comparison of rewarming outcomes among the 3 study groups. Box plot showing the distribution of time to achieve normothermia. The horizontal line within each box represents the median rewarming time. The top and bottom boundaries of the boxes indicate the IQR, and the whiskers extend to the minimum and maximum values. No statistically significant differences were observed (*P*=.460). Control group (group C): forced-air warming at 38 °C; group F1: forced-air warming at 42 °C; group F2: forced-air warming initially set at 42 °C, then reduced to 38 °C upon reaching 36 °C.

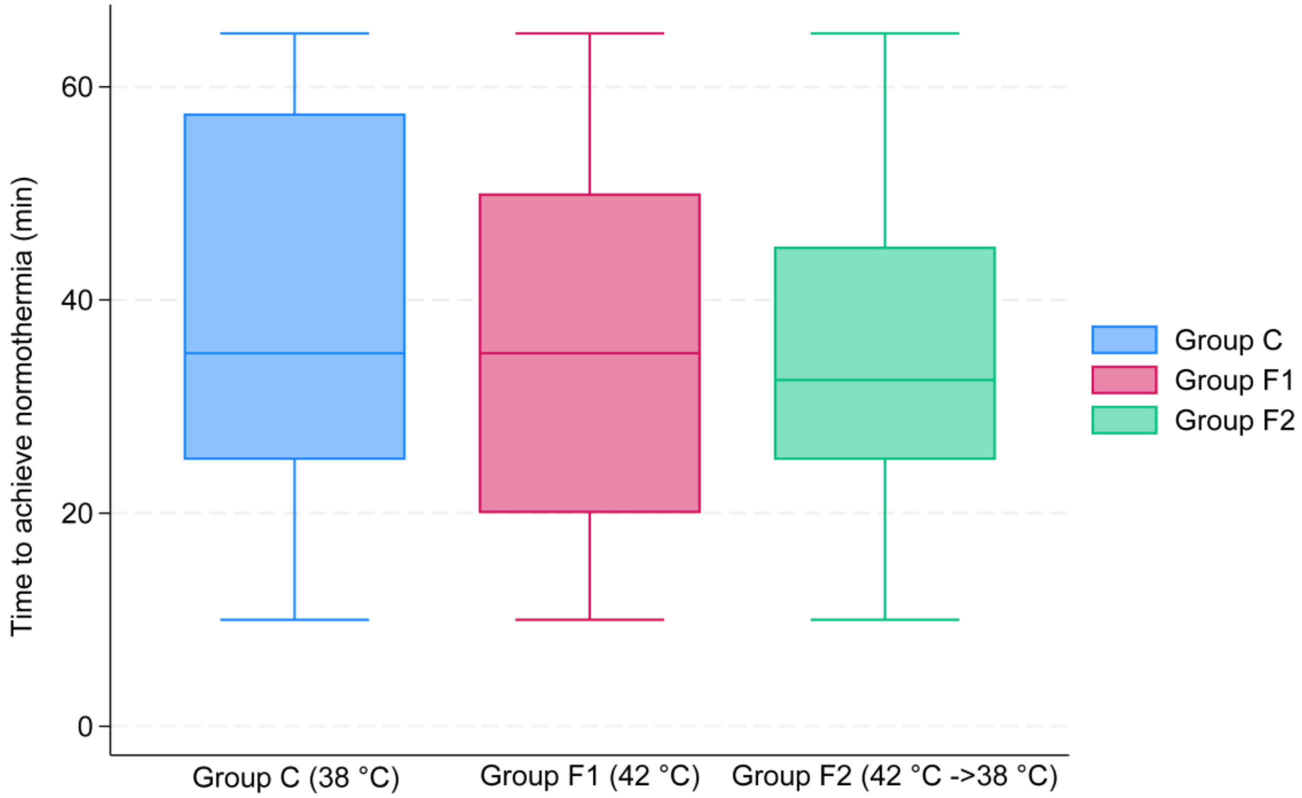
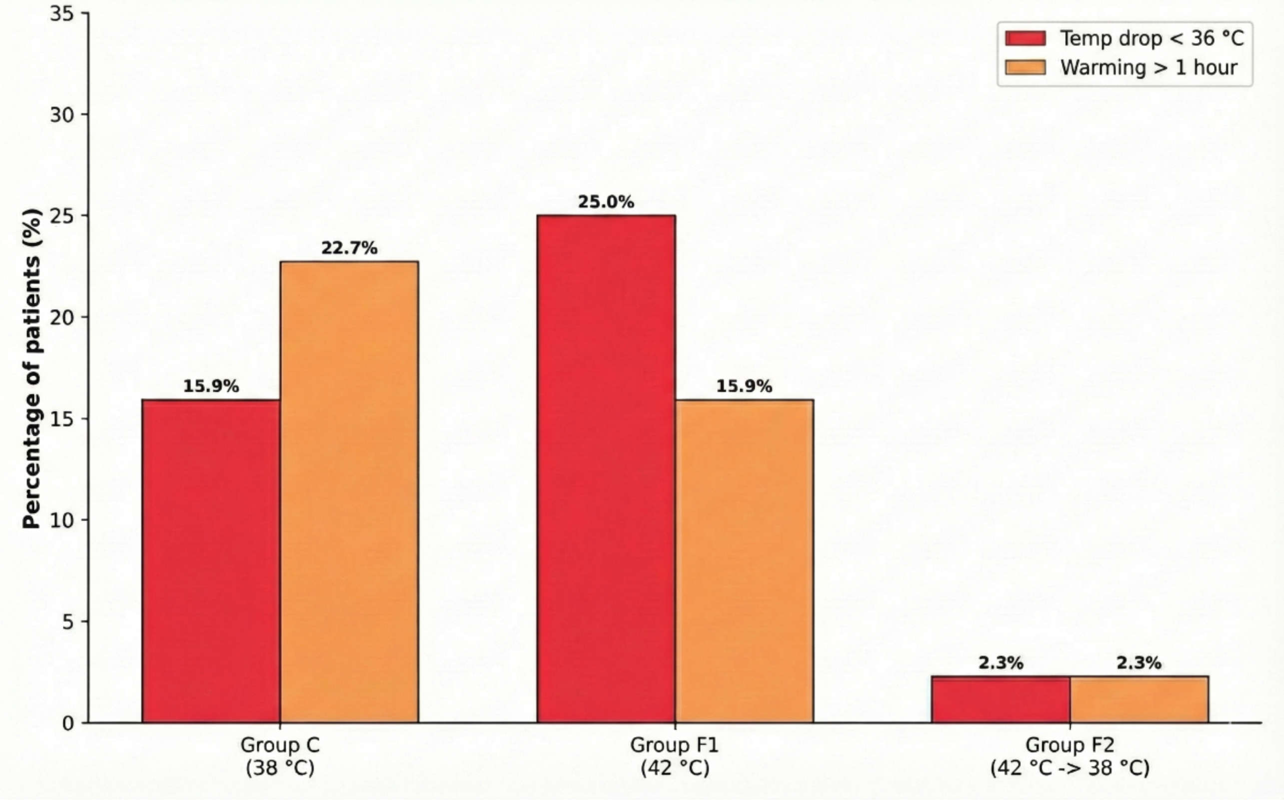


Figure 3. Comparison of rewarming outcomes among the 3 study groups. Clustered bar chart illustrating the incidence of recurrent hypothermia (core temperature dropping <36 °C after achieving normothermia) and the proportion of patients requiring active warming for more than 1 h. Group F2 demonstrated significantly lower rates for both outcomes compared to groups C and F1 (*P*=.009 and *P*=.017, respectively). Control group (group C): forced-air warming at 38 °C; group F1: forced-air warming at 42 °C; group F2: forced-air warming initially set at 42 °C, then reduced to 38 °C upon reaching 36 °C.



Adverse Events

In terms of postoperative adverse events, no significant differences were observed among the 3 groups. Pain was the most frequently reported complication, affecting 22.7% (10/44) of patients in group C and 18.2% (8/44) in both group F1 and group F2 (*P*=.826). Nausea and vomiting occurred infrequently, with an incidence ranging from 2.3%(1/44) to 4.6%(2/44) across the groups (*P*>.999). Shivering was reported in 4.6% (2/44) of patients in group C, 6.8% (3/44) in group F1, and 2.3% (1/44) in group F2 (*P*=.871). Hemodynamic events were rare, comprising 1 case of hypertension in group C, no events in group F1, and 1 case each of hypotension and hypertension in group F2 (all *P*>.999). No patients experienced arrhythmia or other adverse effects. Overall, the incidence of postoperative complications was low and comparable among the groups, supporting the safety of forced-air warming across different temperature settings.

Patient comfort, as measured by satisfaction levels, showed a significant difference among the groups (*P*=.049), along with the average comfort scores (*P*=.039). The proportion of patients reporting being “very much satisfied” was 27.27% (12/44) in group C, 43.18% (19/44) in group F1, and 52.27% (23/44) in group F2 (Table 5). However, there was no significant difference in the thermal comfort scale among the groups (*P*=.131).

Table . Patient satisfaction in the 3 groups.

Parameter	Group C (n=44)	Group F1 (n=44)	Group F2 (n=44)	*P* value
Patient's comfort, n (%)				.049
Very much satisfied	12 (27.27)	19 (43.18)	23 (52.27)	
Somewhat satisfied	29 (65.91)	21 (47.73)	21 (47.73)	
Undecided	2 (4.55)	4 (9.09)	0 (0)	
Not really satisfied	0 (0)	0 (0)	0 (0)	
Not at all satisfied	1 (2.27)	0 (0)	0 (0)	
Patient's comfort (average score), median (IQR)	4 (4-5)	4 (4-5)	5 (5-5)	.039
Thermal comfort scale, n (%)				.131
Hot	1 (2.27)	4 (9.09)	2 (4.55)	
Warm	34 (77.27)	32 (72.73)	31 (70.45)	
Slightly warm	6 (13.64)	2 (4.55)	9 (20.45)	
Neutral	2 (4.55)	6 (13.64)	2 (4.55)	
Slightly cold	1 (2.27)	0 (0)	0 (0)	
Cool	0 (0)	0 (0)	0 (0)	
Cold	0 (0)	0 (0)	0 (0)	
Thermal comfort scale (average score), median (IQR)	2.00(2-2)	2.00(2-2)	2.00 (1.25 - 2)	.676

Discussion

Principal Findings

Postoperative hypothermia is a frequent complication of both general and regional anesthesia, primarily resulting from thermoregulatory impairment and internal heat redistribution [21]. While previous research identified FAW at 42 °C as effective for elderly patients [19], evidence regarding the optimal temperature setting for the general surgical population remains limited. Consequently, this trial aimed to evaluate the most effective and efficient rewarming protocol for patients undergoing various surgical procedures.

In this study, hypothermia was defined as a core temperature <36 °C upon PACU admission. Tympanic thermometry was selected over invasive nasopharyngeal or rectal probes, as used in previous studies [19,22], to prioritize patient comfort during the awake recovery phase.

The principal finding of this study is that increasing the FAW setting from 38 to 42 °C did not yield a statistically significant reduction in the overall rewarming time to normothermia. While group F2 achieved the target temperature approximately 3 to 4 minutes faster than the control group, the precision estimates provided by the 95% CI suggest that this difference is negligible. Given that all groups achieved normothermia within a comparable timeframe, the variation in rewarming speed appears to lack clinical relevance for PACU throughput.

Several physiological factors may explain why higher settings did not produce faster rewarming, a finding that contrasts with some previous studies [19]. Peripheral vasoconstriction can limit the rate of convective heat transfer from the skin to the core, creating a "plateau effect" regardless of the external heat gradient provided by higher FAW settings. Moreover, as the core temperature approaches the normal thermoregulatory threshold, the body initiates vasodilation to redistribute heat, preventing a linear increase in core temperature [21,23]. Device-specific factors, such as automatic safety regulation at higher settings or variability in blanket positioning, may have further minimized the actual difference in heat delivery.

Although rewarming rates were comparable, the group F2 protocol demonstrated superior thermal stability. Unlike group C, which exhibited a significantly higher incidence of prolonged rewarming, the step-down protocol (group F2) effectively minimized the number of "outliers"—patients requiring extended care due to thermal instability. The prolonged rewarming observed in group C is likely due to several physiological and thermodynamic factors. At a lower temperature (38 °C), the gradient between the patient's core temperature and the surrounding warming environment is reduced, leading to a slower rate of heat transfer [24]. Additionally, peripheral vasoconstriction limits blood flow to the skin and extremities, impeding the transport of externally applied heat to the core [25]. Furthermore, the reduced metabolic rate associated with hypothermia decreases endogenous heat generation, collectively contributing to the extended recovery time [21]. Consequently, the requirement for prolonged active warming in groups C and F1 likely contributed to the statistical difference observed in the total duration of PACU stay (*P*=.015). Although the median stay was consistent at 60 minutes across all groups, the distribution of discharge times suggests that while group F2 may not shorten the mandatory minimum recovery time, it optimizes unit throughput by reducing the incidence of prolonged stays.

Clinical Implications

Regarding safety and comfort, the incidence of adverse events—including pain, nausea, vomiting, hemodynamic changes, and shivering—did not differ significantly among the groups, and no severe adverse events were observed. These results align with the safety profile reported by Xu et al [19] However, regarding patient experience, group F2 reported higher satisfaction scores related to comfort during rewarming compared to groups C and F1. This suggests that an initial high-temperature setting effectively enhances thermal comfort, while the subsequent reduction prevents the discomfort associated with overheating.

Intraoperative factors—including ambient cooling, fluid administration, and anesthesia-induced thermoregulatory impairment—are known to significantly impact rewarming. In this study, potential confounding was minimized through a standardized intraoperative care protocol that included routine fluid and forced-air warming. Furthermore, randomization successfully balanced these physiological stressors across study arms; as shown in Table 2, there were no significant differences in operative duration, total fluid volume, or anesthetic technique. Consequently, the observed differences in PACU outcomes can be primarily attributed to the specific postoperative warming protocols rather than intraoperative disparities.

Strengths and Limitations

This study has notable strengths and limitations. A key strength is the rigorous randomization and standardized intraoperative care, which successfully balanced potential confounders such as operative duration and fluid volume across study arms. However, several limitations exist. First, strict environmental control of the PACU was challenging due to the open-plan nature of the unit. The ambient temperature fluctuated between 22 and 24 °C, which serves as a potential environmental confounder influencing convective heat loss. Nevertheless, this variation reflects real-world clinical conditions, potentially enhancing the ecological validity of our results. Second, regarding measurement, reliance on tympanic thermometry introduces inherent variability compared to the gold standard of invasive core monitoring. We acknowledge that readings can be affected by factors such as probe positioning, cerumen obstruction, and post-anesthetic peripheral vasoconstriction. To mitigate these inaccuracies, we strictly standardized the technique by using the same device and assessing the ipsilateral ear throughout the study, aiming to balance measurement precision with patient comfort in the awake state. Third, data collection involved intermittent recordings at 5-minute intervals rather than continuous electronic monitoring. While this frequency is clinically practical, it may lack the temporal resolution to capture rapid, transient temperature fluctuations during the active rewarming phase, potentially masking the true extent of thermal variability. Finally, we did not perform a formal cost-effectiveness analysis. Although the step-down protocol (group F2) showed potential for optimizing PACU throughput, future studies including economic evaluations are needed to confirm the financial implications of these warming strategies.

Conclusions

In summary, although varying temperature settings of forced-air warming systems produced comparable rewarming times, the protocol involving an initial setting of 42 °C followed by a reduction to 38 °C (group F2) was associated with superior maintenance of normothermia and a significantly lower incidence of postoperative hypothermia recurrence. These findings underscore the potential benefits of implementing optimized warming protocols to enhance patient outcomes in the PACU.

Acknowledgments

The authors would like to thank all patients who willingly participated in this study, without whom the research would not have been possible. We gratefully acknowledge the contribution of Ms. Rojnarin Komonhirun, M.Ed., for her valuable assistance with data analysis. We also thank the PACU staff of Ramathibodi Hospital for their support in facilitating the data collection process.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

Due to privacy concerns, the datasets generated or analyzed during this study are not publicly available. However, deidentified data may be provided by the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: KP (lead), SS (equal), RS (supporting), WW (supporting)
Data curation: KP (lead), SS (equal)
Formal analysis: KP (lead), SS (equal)
Investigation: KP (lead), SS (equal)
Methodology: KP (lead), SS (equal)

Resources: KP (lead), SS (equal), WW (supporting)
Supervision: KP (lead), RS (supporting)
Validation: KP (lead), SS (equal)
Visualization: KP (lead), SS (equal), RS (supporting), WW (supporting)
Writing – original draft: KP (lead), SS (equal)
Writing – review & editing: KP (lead), SS (equal), RS (supporting), WW (supporting)

Conflicts of Interest

None declared.

References

1. Shenoy L, Krishna HM, Kalyan N, Prasad KH. A prospective comparative study between prewarming and cowarming to prevent intraoperative hypothermia. J Anaesthesiol Clin Pharmacol 2019;35(2):231-235. [doi: 10.4103/joacp.JOACP_353_17] [Medline: 31303714]
2. Ashoobi MT, Shakiba M, Keshavarzmotamed A, Ashraf A. Prevalence of postoperative hypothermia in the post-anesthesia care unit. Anesth Pain Med 2023 Oct;13(5):e136730. [doi: 10.5812/aapm-136730] [Medline: 38028114]
3. Torossian A, Bräuer A, Höcker J, Bein B, Wulf H, Horn EP. Preventing inadvertent perioperative hypothermia. Dtsch Arztebl Int 2015 Mar 6;112(10):166-172. [doi: 10.3238/arztebl.2015.0166] [Medline: 25837741]
4. Shaw CA, Steelman VM, DeBerg J, Schweizer ML. Effectiveness of active and passive warming for the prevention of inadvertent hypothermia in patients receiving neuraxial anesthesia: a systematic review and meta-analysis of randomized controlled trials. J Clin Anesth 2017 May;38:93-104. [doi: 10.1016/j.jclinane.2017.01.005] [Medline: 28372696]
5. Rauch S, Miller C, Bräuer A, Wallner B, Bock M, Paal P. Perioperative hypothermia - a narrative review. Int J Environ Res Public Health 2021 Aug 19;18(16):8749. [doi: 10.3390/ijerph18168749] [Medline: 34444504]
6. Munday J, Duff J, Wood FM, Sturgess D, Ralph N, Ramis MA. Perioperative hypothermia prevention: development of simple principles and practice recommendations using a multidisciplinary consensus-based approach. BMJ Open 2023 Nov 14;13(11):e077472. [doi: 10.1136/bmjopen-2023-077472] [Medline: 37963694]
7. Grant MC, Engelman DT. Enhanced recovery after surgery: overarching themes of the ERAS® Society Guidelines & Consensus Statements for Adult Specialty Surgery. Perioper Med (Lond) 2025;14(1):120. [doi: 10.1186/s13741-025-00590-0] [Medline: 41168801]
8. Wongsripuemtet P, Leelanokrom R, Tangwiwat S, Vorapaluk P, Karnjanarachata C, Karnjanawanichkul O, et al. Clinical guidelines for prevention and management of perioperative hypothermia by the Royal College of Anesthesiologists of Thailand. Thai J Anesthesiol 2024;50:287-300 [FREE Full text]
9. Lynch S, Dixon J, Leary D. Reducing the risk of unplanned perioperative hypothermia. AORN J 2010 Nov;92(5):553-562. [doi: 10.1016/j.aorn.2010.06.015] [Medline: 21040819]
10. Bräuer A, Quintel M. Forced-air warming: technology, physical background and practical aspects. Curr Opin Anaesthesiol 2009 Dec;22(6):769-774. [doi: 10.1097/ACO.0b013e328331d134] [Medline: 19734783]
11. Warttig S, Alderson P, Campbell G, Smith AF. Interventions for treating inadvertent postoperative hypothermia. Cochrane Database Syst Rev 2014 Nov 20;2014(11):CD009892. [doi: 10.1002/14651858.CD009892.pub2] [Medline: 25411963]
12. Pruitthitada S, Sinchai S. Effective of preoperative force air warming on perioperative hypothermia in patients undergoing caesarean section under spinal anesthesia. Thai J Anesthesiol 2022;48(2):87-98 [FREE Full text]
13. Becerra Á, Valencia L, Villar J, Rodríguez-Pérez A. Short-periods of pre-warming in laparoscopic surgery. A non-randomized clinical trial evaluating current clinical practice. J Clin Med 2021 Mar 3;10(5):1047. [doi: 10.3390/jcm10051047] [Medline: 33802512]
14. Jun JH, Chung MH, Kim EM, et al. Effect of pre-warming on perioperative hypothermia during holmium laser enucleation of the prostate under spinal anesthesia: a prospective randomized controlled trial. BMC Anesthesiol 2018 Dec 22;18(1):201. [doi: 10.1186/s12871-018-0668-4] [Medline: 30579334]
15. Yoo JH, Ok SY, Kim SH, et al. Effects of 10-min of pre-warming on inadvertent perioperative hypothermia in intraoperative warming patients: a randomized controlled trial. Anesth Pain Med (Seoul) 2020 Jul 31;15(3):356-364. [doi: 10.17085/apm.20027] [Medline: 33329836]
16. Parameeanon M, Poomepraman W, Sarachai W. Effect of pre-warming by forced air warmer during total hip and knee arthroplasties for perioperative and postoperative hypothermia. Thai J Anesthesiol 2020;46(2):88-95 [FREE Full text]
17. Zhang R, Chen X, Xiao Y. The effects of a forced-air warming system plus electric blanket for elderly patients undergoing transurethral resection of the prostate: a randomized controlled trial. Medicine (Baltimore) 2018 Nov;97(45):e13119. [doi: 10.1097/MD.0000000000013119] [Medline: 30407328]
18. Tyvold SS. Preventing hypothermia in outpatient plastic surgery by self-warming or forced-air-warming blanket: a randomised controlled trial. Eur J Anaesthesiol 2019 Nov;36(11):843-850. [doi: 10.1097/EJA.0000000000001087] [Medline: 31567576]
19. Xu H, Xu G, Ren C, Liu L, Wei L. Effect of forced-air warming system in prevention of postoperative hypothermia in elderly patients: a prospective controlled trial. Medicine (Baltimore) 2019;98(22):e15895. [doi: 10.1097/MD.0000000000015895]

20. Pisitsak C, Virankabutra T, Deewong K, Pornprasertsuk M. Postoperative hypothermia: incidence and associated factors. Thai J Anesthesiol 2011;37:93-103.
21. Sessler DI. Temperature monitoring and perioperative thermoregulation. Anesthesiology 2008 Aug;109(2):318-338. [doi: 10.1097/ALN.0b013e31817f6d76] [Medline: 18648241]
22. Liu X, Shi Y, Ren C, Li X, Zhang Z. Effect of an electric blanket plus a forced-air warming system for children with postoperative hypothermia: a randomized controlled trial. Medicine (Baltimore) 2017;96(26):e7389. [doi: 10.1097/MD.0000000000007389]
23. Sessler DI. Complications and treatment of mild hypothermia. Anesthesiology 2001 Aug;95(2):531-543. [doi: 10.1097/00000542-200108000-00040] [Medline: 11506130]
24. Taylor EE, Carroll JP, Lovitt MA, et al. Active intravascular rewarming for hypothermia associated with traumatic injury: early experience with a new technique. Proc (Bayl Univ Med Cent) 2008 Apr;21(2):120-126. [doi: 10.1080/08998280.2008.11928375] [Medline: 18382749]
25. Ereth MH, Lennon RL, Sessler DI. Limited heat transfer between thermal compartments during rewarming in vasoconstricted patients. Aviat Space Environ Med 1992 Dec;63(12):1065-1069. [Medline: 1456917]

Abbreviations

ASA: American Society of Anesthesiologists
FAW: forced-air warming
PACU: postanesthesia care unit

Edited by J Shiffermiller; submitted 10.Oct.2025; peer-reviewed by P Sawaddiruk; revised version received 10.Dec.2025; accepted 25.Dec.2025; published 28.Jan.2026.

Please cite as:
Pasutharnchat K, Seangrung R, Sirisophaphong S, Wongkum W
Forced-Air Warming Temperature Settings for Treating Postoperative Hypothermia in the Postanesthesia Care Unit: Randomized Controlled Trial
JMIR Perioper Med 2026;9:e85045
URL: https://periop.jmir.org/2026/1/e85045
doi:10.2196/85045

Virtual Reality for the Management of Postoperative Pain and Anxiety in Children and Adolescents Undergoing Nuss Repair of Pectus Excavatum: Randomized Controlled Trial

Charlotte M Walter[1], MD; Dillon Froass[2], BS; Nora Bell[3], MD; Lauren Haack[4], MD; Chloe Boehmer[1], MA; Claudia Bruguera Torres[1], MD; Rachel Spivak[1], BA; Max Chou[5], MD; Kristie Geisler[1], BS; Keith O'Conor[6], MD; Sara E Williams[7], PhD; Lili Ding[1], PhD; Christopher D King[1], PhD; Vanessa A Olbrecht[8,9], MBA, MD

[1]Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, United States

[2]The Ohio State University Wexner Medical Center, Columbus, OH, United States

[3]Department of Pediatrics, Vanderbilt University, Nashville, TN, United States

[4]Department of Surgery, Wayne State University, Detroit, MI, United States

[5]Department of Anesthesiology, University of Cincinnati Medical Center, Cincinnati, OH, United States

[6]Department of Anesthesiology, Penn State Milton S. Hershey Medical Center, Hershey, PA, United States

[7]Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University, Palo Alto, CA, United States

[8]Department of Anesthesiology, Nemours Children's Hospital, Wilmington, DE, United States

[9]Sidney Kimmel Medical College, Thomas Jefferson University Hospital, Philadelphia, PA, United States

Corresponding Author:
Charlotte M Walter, MD
Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, United States

Abstract

Background: Virtual reality (VR) is a novel technology with implications for pain and sensory processing. VR may serve as a novel, scalable method to deliver clinically validated therapy for pain management as an alternative or adjunct to opioids for acute pain. Given that psychological factors and pain perception are both components of postoperative pain, it may also be beneficial to incorporate modalities that decrease anxiety, such as active relaxation and guided meditation with VR. Unfortunately, these therapies are not widely available due to multiple barriers. VR has the potential to deliver pain-reducing, psychologically based therapy to children, thereby enhancing multimodal analgesia and potentially decreasing opioid use. This study investigates the role of VR in reducing pain and anxiety after surgery. Given the substantial risks associated with opioid use, particularly in younger populations, alternative pain management strategies are crucial.

Objective: The primary aim of this study was to evaluate the efficacy of VR as a nonpharmacological intervention for managing postoperative pain intensity, pain unpleasantness, anxiety, and opioid use in children and adolescents undergoing Nuss repair of pectus excavatum.

Methods: A single-center, prospective, randomized, controlled trial was conducted at a tertiary care children's hospital and research center. Ninety children and adolescents (8-18 y) undergoing the Nuss procedure were randomized to guided relaxation or mindfulness VR (n=30) and distraction-based gaming VR (n=30), combined to form the VR group (n=60), and a control group using a passive 360° video (n=30). Patients received a 10-minute session on postoperative days 1 and 2. Pain intensity, pain unpleasantness, and anxiety were evaluated before and 0-, 15-, and 30-minute post-session. In-hospital pain scores, anxiety scores, and opioid use were collected.

Results: Children and adolescents who participated in VR reported a significantly greater decrease in pain intensity from baseline (0.41, SE 0.23) compared with those in the 360° video group at 30 minutes (P=.04) before multiplicity adjustment but not after multiplicity adjustment. There were no significant differences in pain scores or opioid use between the VR and control groups on postoperative day 1 or 2, nor were there changes in pain unpleasantness or anxiety at any time after the intervention.

Conclusions: Daily, 10-minute VR sessions provided some trends toward transient analgesic and anxiolytic effects, albeit none that were statistically significant. VR did not significantly decrease overall pain scores or opioid usage, possibly due to the limited intervention duration and high standardized opioid use. Future studies should investigate extended and more frequent VR sessions and the integration of VR with other therapeutic modalities.

Trial Registration: ClinicalTrials.gov NCT04351776; https://clinicaltrials.gov/study/NCT04351776

(JMIR Perioper Med 2026;9:e80902) doi:10.2196/80902

KEYWORDS

analgesia; anxiety; distraction-based virtual reality; pediatric anesthesia; prospective studies; child; human; virtual reality; pediatric pain medicine; acute postoperative pain

Introduction

Background

Multimodal pain management techniques for acute postoperative pain are commonly studied and utilized [1]; opioids continue to be the cornerstone of postoperative pain management. Opioid misuse continues to be a major public health issue in the United States, with children and adolescents particularly vulnerable, as many are initially exposed to opioids prescribed for pain management [2-4]. Furthermore, the risk of future opioid overdose significantly increases with the quantity of pills prescribed; adolescents receiving 30 or more pills have a 35% higher rate of overdose than those prescribed 18 or fewer pills [5]. The prescription of opioid analgesics is a well-documented pathway to misuse, opioid use disorder, and overdose [6].

The Nuss procedure, performed to repair pectus excavatum, is associated with severe postoperative pain [7]. Effective management of postoperative pain after this surgery is crucial, as alleviating pain can enhance patient satisfaction and reduce complication rates [8]. Effective pain management techniques and regimens vary across pediatric institutions and have begun including intercostal nerve cryoablation [8,9]. Opioid use during recovery from the Nuss procedure is common, with one study finding that patients used opioids for a median of 8 days with an IQR of 6 - 10 days [10]. Given the absence of standardized postoperative pain management protocols and the high usage of opioids following the Nuss procedure, it is essential to explore nonpharmacologic pain control adjuncts for these patients.

Virtual reality (VR) technology provides an immersive, multisensory, and 3D environment that enables individuals to experience a modified reality, creating a sense of “presence” for each individual [11]. There is a clear need for alternative pain management methods, including nonpharmacologic techniques. VR has been shown to be effective in reducing perioperative and postoperative anxiety in pediatric patients. Studies show significant reductions in anxiety in pediatric patients immediately after distraction-based gaming virtual reality (VR-D) sessions, with some effects lasting for at least 15 minutes post-intervention [12,13]. Two approaches—VR-D and guided relaxation–based virtual reality (VR-GR)—are being researched for their effectiveness in reducing pain and anxiety following surgery.

VR-D immerses patients in engaging experiences that help divert attention from pain or anxiety, providing effective short-term relief. Integration of these techniques is challenging in the perioperative space, with limited providers and resources and high costs limiting its feasibility. VR can be used anywhere, anytime, with access to a headset.

Gate control theory suggests that distraction can be a valuable tool for pain management, as attentional load is fixed, and distraction toward a pleasant experience means less attention to pain [14,15]. It has been associated with immediate and short-term reductions in postoperative pain intensity and unpleasantness. VR-D techniques have been shown to decrease acute pain in children and adults [16-18]. Single sessions of VR-D have been shown to reduce postoperative pain for up to 30 minutes in some cases, regardless of baseline pain catastrophizing levels, suggesting broad applicability across pediatric populations experiencing postoperative pain [19]. The use of VR-D has also demonstrated pain reduction comparable to opioid use in burn injury patients during wound cleaning [20]. While VR-D is particularly effective for short-term pain management, additional strategies like guided relaxation may be needed for longer-lasting pain relief [21].

VR-GR seeks to provide more sustained pain relief by combining distraction with mind-body techniques, such as guided relaxation or mindfulness within the VR environment. Psychological factors—including calmness, fear, anxiety, and depression—affect the subjective experience of pain [22]. Resilience has been negatively associated with pain unpleasantness, potentially serving as a protective factor in patients with higher baseline anxiety [22,23]. Incorporating active relaxation and guided meditation techniques may significantly contribute to pain reduction. This combination of settling the mind to increase resilience and distraction from acute pain may play a significant role in acute pain reduction [21]. VR-GR may further improve anxiety reduction, especially in children with higher anxiety sensitivity [12,24]. Although VR-GR may offer additional benefits for sustained pain relief compared to distraction alone, its effects were also primarily transient [21]. Using VR to perform guided relaxation could expand the benefits of these nonpharmacological pain management techniques to more children, including those having surgery.

Overall, VR is a promising nonpharmacologic tool for managing postoperative pain and anxiety in children and adolescents. It can potentially enhance the perioperative experience, reduce reliance on pharmacological interventions, and increase patient and family satisfaction. However, randomized controlled trials are needed to establish standardized protocols and explore VR integration with other therapies, such as biofeedback, for more durable outcomes [13,21,25].

Aim

In this prospective, randomized, controlled clinical trial, we compare the short-term efficacy of immersive VR in decreasing acute postoperative pain (primary outcome), anxiety, and opioid consumption following pectus excavatum repair. We hypothesize that the use of VR will be more effective at reducing pain, anxiety, and opioid use as compared to the control group in this patient population.

Methods

The original protocol for this study has been published [26].

Study Design and Setting

This single-center, randomized, unblinded clinical trial was conducted at Cincinnati Children's Hospital Medical Center (CCHMC), a 670-bed tertiary care academic children's hospital. The recruitment began on July 10, 2020, and the study was completed on July 30, 2023. The COVID-19 pandemic delayed study completion. We recruited children and adolescents undergoing Nuss repair of pectus excavatum to investigate the role of VR in the management of postoperative pain and anxiety.

Ethical Considerations

This study complies with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement [27] and the Consolidated Standard of Reporting Trials (CONSORT) statement [28]. The CCHMC Institutional Review Board approved this study (IRB 2019 - 1090) on November 26, 2019, and it was conducted per the rules and regulations for ethical research. This study was registered at ClinicalTrials.gov on April 3, 2020 (NCT04351776). Written informed parental consent and patient assent (children >11 years) were obtained from all participants before enrollment into this study. Patients received a small stipend for participation. All identifying patient information was kept private and confidential.

Patients and Recruitment

Patients

This study recruited 90 patients (30 patients per group), ages 8 to 18 years, undergoing Nuss repair of pectus excavatum surgery. Informed parental consent and patient assent were obtained before enrollment into this study.

The inclusion criteria were as follows: patients were (1) between the ages of 8 and 18 years, (2) able to read, understand, and speak English, (3) presenting for Nuss repair of pectus excavatum, and (4) followed by the acute pain service following surgery.

The exclusion criteria were as follows: patients with (1) a history of developmental delay, uncontrolled psychiatric conditions, or neurological conditions, (2) a history of seizures, epilepsy, vertigo, or significant motion sickness/nausea/vomiting, or (3) any condition that would preclude the application of the VR headset, such as craniofacial abnormalities.

Recruitment

Approximately 150 Nuss repair surgeries are performed at CCHMC each year. Therefore, our recruitment target of 90 patients was well within achievable limits. During the study, patients who underwent Nuss repair were recruited continuously until we met the targeted enrollment. The operating room schedule and surgical patient list were reviewed for potentially eligible patients, who were approached for recruitment before surgery. If patients wished to participate, consent (and assent for patients >11 years of age) was obtained, and eligibility criteria were verified. We recruited about 2 patients per week. Recruiting stopped during the COVID-19 pandemic, when elective surgeries were not performed, delaying study completion.

Randomization

Potential patients were identified using the operating room schedule and the pectus surgery list provided by the surgery team. Eligible participants were randomized (1:1:1) into three groups: active distraction-based guided relaxation virtual reality (VR-DGR, n=30) and active VR-D (n=30)—collectively the VR group (n=60)—and a control group—passive 360° video (360-V) without instructions, sound, guided relaxation, or active patient involvement (n=30).

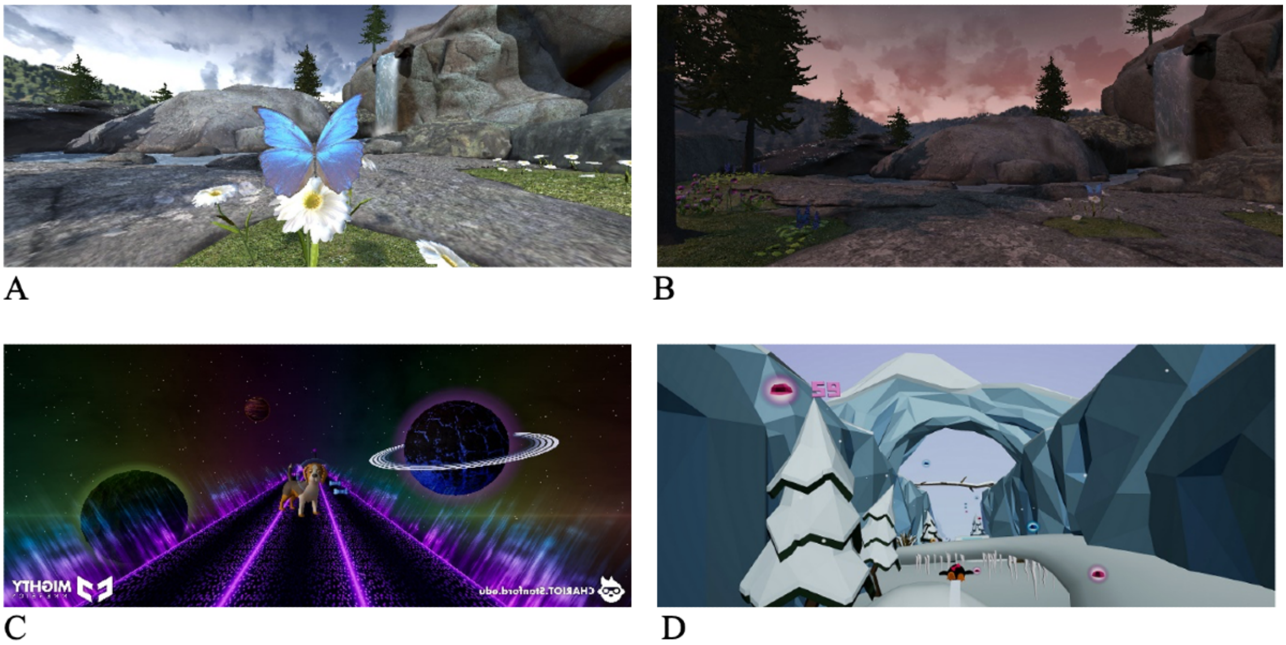
VR Technology

All participants used a Starlight Xperience VR all-in-one device and software developed specifically for hospital settings. This technology is a customized version of the Lenovo Mirage Solo with a Daydream VR headset. It is easy to disinfect to comply with hospital infection safety protocols. Importantly, an integrated headphone device provides audio content, and the patients use head movements and a handheld controller for interaction and navigation. It is commercially available (not Food and Drug Administration–regulated) and was supplied by the Starlight Children's Foundation.

VR-DGR and 360-V participants used the Mindful Aurora application, developed by the Stanford University Childhood Anxiety Reduction through Innovation and Technology program, to deliver relaxation/mindfulness content, which presents a relaxing nature scene with prompts instructing patients to actively slow and pace their breathing in conjunction with the movement of objects in the VR environment. 360°-V participants experienced the same relaxing nature scene without guided relaxation prompts; the 360°-V group also did not receive any audio and thus did not experience an immersive environment.

VR-D participants had the option to choose and play one of the three games: Space Pups, in which the participant controls a puppy in space and collects treats to music; Pebbles the Penguin, in which the participant controls a penguin sliding down a mountain to collect pebbles; or Wonderglade, in which the participant can play five different mini-carnival games (Figure 1).

Figure 1. Scenes from the Mindful Aurora application used in distraction-based guided relaxation virtual reality (VR-DGR) and 360° video (360-V) (A and B), and scenes from Space Pups (C) and Pebbles the Penguin (D) used in distraction-based gaming virtual reality (VR-D).



Procedure

Consent and assent were obtained before the visit. Patient characteristics, demographics, weight, and pain scores were collected preoperatively. All patients enrolled in this study received standard postoperative care via the CCHMC Pectus Surgery Pain Management Protocol, which standardizes all medications received by all pectus patients. This includes non-opioid pain medication such as pregabalin, acetaminophen, ketorolac, methocarbamol, and diazepam. All participants received the same non-opioid medications. Before the first session, patients completed the Childhood Anxiety Sensitivity Index to establish baseline anxiety levels and the Pain Catastrophizing Scale (PCS) for children. Patients were visited daily for one 10-minute session. Every effort was made to ensure the consistent timing of the visits for all patients. Sessions were completed beginning on postoperative day (POD) 1, then daily until the day of discharge, or until POD 3.

Patients were trained to use the technology before the first VR session. All participants received a device tutorial that taught them how to use the device and introduced them to the VR software. Patients received a script about VR-GR, VR-D, or 360-V, depending on the group to which they are assigned. During each session, patients completed a 10-minute session of either VR-GR, VR-D, or 360-V, per assigned group. Patients were asked to rate their pain intensity, pain unpleasantness, and anxiety via Numerical Rating Scale (NRS), before, immediately after, and 15 and 30 minutes after each session. Pain and anxiety scores and opioid use/day were recorded in REDCap (Research Electronic Data Capture).

Data Collection

The primary outcome measure was pain intensity (NRS), measured before, immediately after, and 15 and 30 minutes following each session on POD1 and POD2. Secondary outcomes included opioid use on POD0, POD1, and POD2 and pain area under the curve (AUC) on POD1 and POD2, pain unpleasantness, and anxiety scores before and 0, 15, and 30 minutes after each session on POD1 and POD2 to establish change from baseline.

For each eligible participant, data were collected from their patient history/interview and the electronic medical record in a standardized case report form in the REDCap system. Inpatient opioid use was identified from the patient's electronic medical record based on documentation in the medication administration record and transferred to REDCap. All opioid quantities were translated to morphine milligram equivalents (MME) and summed to determine total morphine equivalents per 24-hour period (midnight-to-midnight) during the patients' inpatient stay. Measures used in the study are summarized in Table 1.

Table . Scales and questionnaires used in the study.

Scales and questionnaires	Definition
Pain intensity and pain unpleasantness	
Numerical Rating Scale (NRS) [29]	The NRS is the most common validated self-report measure of pain intensity and pain unpleasantness. It involves verbally asking for an estimate of pain intensity using numbers from 0 (no pain) to 10 (maximal pain). Pain was described as being like listening to music; pain intensity is the volume, and pain unpleasantness is how much the music is disliked [30]. It requires no equipment to administer or score.
Pain intensity across all postoperative days	Area under the time-pain score (NRS) curve using the trapezoidal rule (pain AUC[a]) measured pain intensity across all postoperative days 1 and 2.
Anxiety	
Pain Catastrophizing Scale for Children (PCS-C) [31]	A validated 13-item questionnaire (each rated on a 5-point scale, 0-4) designed to measure pain catastrophizing in children of age 8 - 17 years. It is adapted from the adult version and assesses three key aspects of pain-related negative thinking: rumination, magnification, and helplessness.
Child Anxiety Sensitivity Index (CASI) [32]	A validated 18-item survey that measures perceived anxiety symptoms. Participants respond to each item on a 3-point scale (eg, “none,” “some,” and “a lot”). The total score is calculated by summing the responses, with higher scores indicating greater anxiety sensitivity. The total scores range from 18 to 54. CASI has been used in VR[b] studies in adolescents of age 10 - 21 years [33].
Numerical Rating Scale-Anxiety (NRS-A) [34]	A validated self-report numeric 0 - 10 anxiety scale that is easy to administer to children. The NRS-A is easy to administer and can be used quickly to assess anxiety levels.
Opioid use	
NIH morphine milligram equivalents (MME) per day [35,36]	Standardizes a metric for quantifying and comparing doses of different opioids. However, MMEs serve as a common metric for comparing different opioids.

[a]AUC: area under the curve.

[b]VR: virtual reality.

Statistical Analysis

Sample Size Calculation

Sample size was based on the feasibility of conducting this clinical study and unpublished preliminary data that assessed the impact of a single VR-D session on pain intensity in children and adolescents after surgery, with a goal of 80% power to detect differences in mean changes of 1 between VR and 360-V (given pilot data which showed average change in pain intensity of −1 [SD 1.2] and correlation of 0.88). Assuming similar results in the passive control group, a sample size of 30 per group will have 80% power to detect differences in mean changes of 1 between VR-GR and the two control groups. Significance (α) is .025 to control for 2 comparisons.

The VR-DGR and VR-D groups were combined into a single VR group for data analysis because both groups utilized active, distraction-based, immersive VR experiences. VR-DGR did not provide participants with feedback on their respiratory or heart rates. Consequently, it functioned as a distraction-based technique and did not significantly differ from the VR-D experience. Therefore, we combined the two groups, as both were fundamentally distraction-based.

A sample size of 60 for the treatment group and 30 for the control group will have 80% power to detect differences in mean changes of 1 between VR and control.

Data Analysis

All statistical analyses were performed using SAS 9.4 (SAS Institute). Patient demographics were described using mean (SD) or median (IQR) for continuous variables, depending on data distribution, and frequency (percentage) for categorical variables and compared between groups using *t* tests, Wilcoxon rank-sum tests, chi-square tests, or Fisher exact tests, as appropriate. Pain AUCs on POD1 and POD2 were calculated as the area under the time-pain score curve using the trapezoidal rule. MMEs, a metric for quantifying and comparing doses of different opioids, were derived for POD0, POD1, and POD2. Change from baseline on pain intensity, pain unpleasantness, and anxiety immediately after and 15 and 30 minutes following each session was calculated as the postinterval value minus the baseline (before session) value on POD 1 and POD2. Mixed effects models for repeated measures were used for pain AUC, MME, and change from baseline on pain intensity, pain unpleasantness, and anxiety outcomes. All mixed effects models included the intervention group and POD as fixed effects and the participant as a random effect. Models for the change from baseline outcomes also included baseline value, time (0-, 15-, and 30-min post-intervention), and group and time interaction as fixed effects. Missing data in the outcomes were examined for pattern and assumed missing at random and handled using full information maximum likelihood (FIML) for mixed effects models. Sidak adjustment for multiplicity was used for change from baseline in pain intensity between intervention groups at 3 time points (immediately after and 15 and 30 minutes following each session).

Results

Participants

Ninety patients were enrolled in the study (60 in VR and 30 in 360-V; Figure 2). The participants comprised 73 male and 17 female patients and had an American Society of Anesthesiologists (ASA) Physical Status Classification System score of 1 - 3, with a mean age of 15.5 (SD 1.4) years. The 2 groups had no differences demographically except for a difference in PCS scores (VR: median 18, IQR 15 - 23; VR-360: median 24, IQR 16 - 28; *P*=.04). Patients were primarily male, adolescent, and Caucasian. This is consistent with the demographics of patients with pectus excavatum [37], and these are the patients most likely to undergo the Nuss procedure [38] (Table 2). All patients had at least 1 observation on all repeated-measured outcomes (pain AUC, MME, pain intensity, pain unpleasantness, and anxiety), and all available data were included in the mixed effects models for the outcomes.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) 2010 flow diagram. VR: virtual reality.

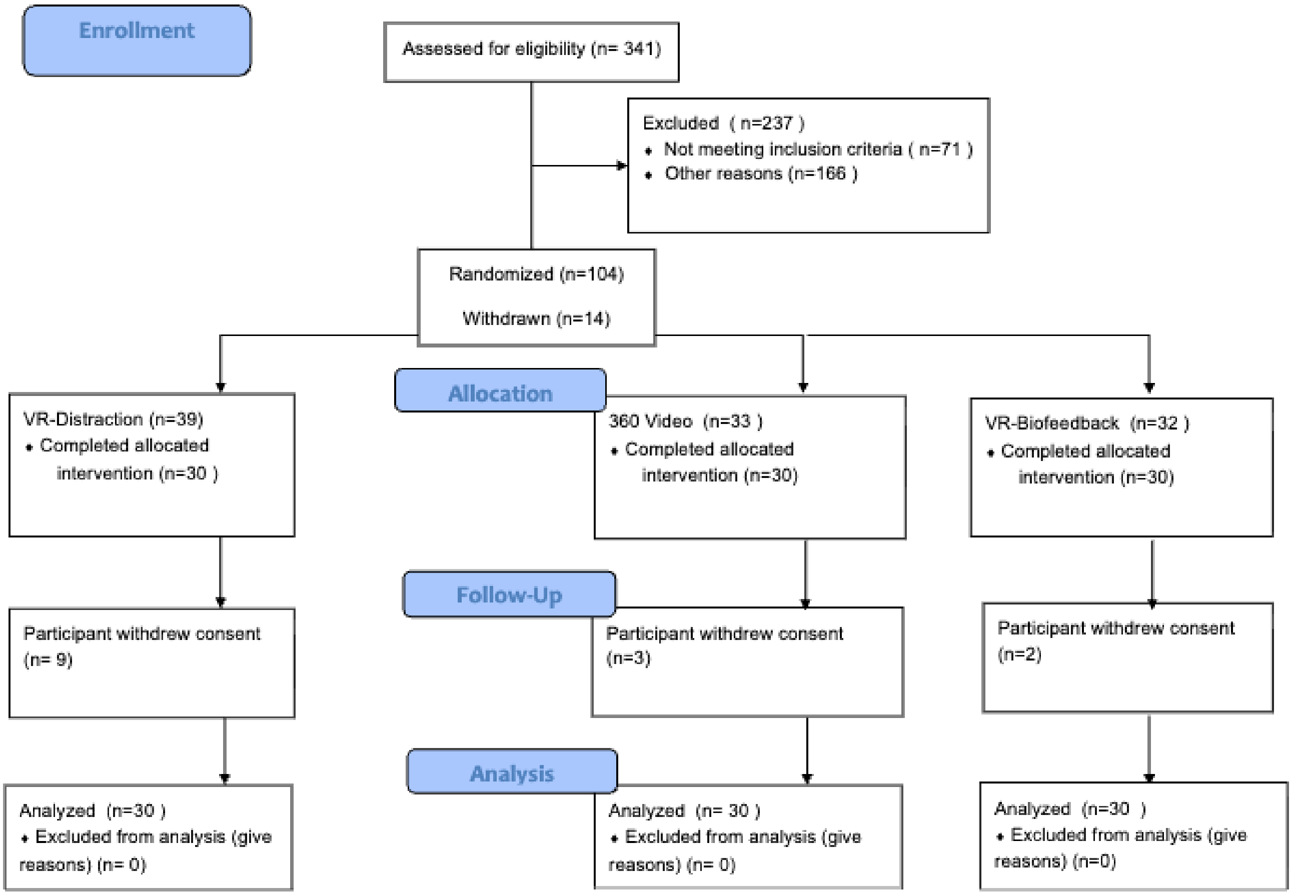


Table . Patient characteristics.

Characteristic	VR[a]	360° video	Overall	*P* value (test)
Age (y), mean (SD)	15.6 (1.4)	15.1 (1.5)	15.5 (1.4)	.10
ASA[b] physical status, n (%)				.34
I	3 (5)	4 (13.3)	7 (7.8)	
II	44 (73.3)	21 (70.0)	65 (72.2)	
III	13 (21.7)	5 (16.7)	18 (20.0)	
Race, n (%)				.55
Caucasian	57 (95)	30 (100)	87 (96.7)	
African American	0 (0)	0 (0)	0 (0)	
Asian	0 (0)	0 (0)	0 (0)	
Other	3 (5)	0 (0)	3 (3.3)	
Ethnicity, n (%)				.25
Hispanic	3 (5)	0 (0)	3 (3.3)	
Non-Hispanic	56 (93.3)	28 (93.3)	94 (93.3)	
Unknown	1 (1.7)	2 (6.7)	3 (3.3)	
Sex, n (%)				.70
Male	48 (80)	25 (83.3)	73 (81.1)	
Female	12 (20)	5 (16.7)	17 (18.9)	
Weight (kg), mean (SD)	58.9 (9.8)	56.3 (10)	58.1 (9.9)	.24
CASI[c] score, mean (SD)	29.3 (5)	30.9 (4.8)	29.8 (5)	.14
Pain Catastrophizing Scale (PCS), median (IQR)	18 (15-23)	24 (16-28)	19 (15-26)	.04

[a]VR: virtual reality.

[b]ASA: American Society of Anesthesiologists.

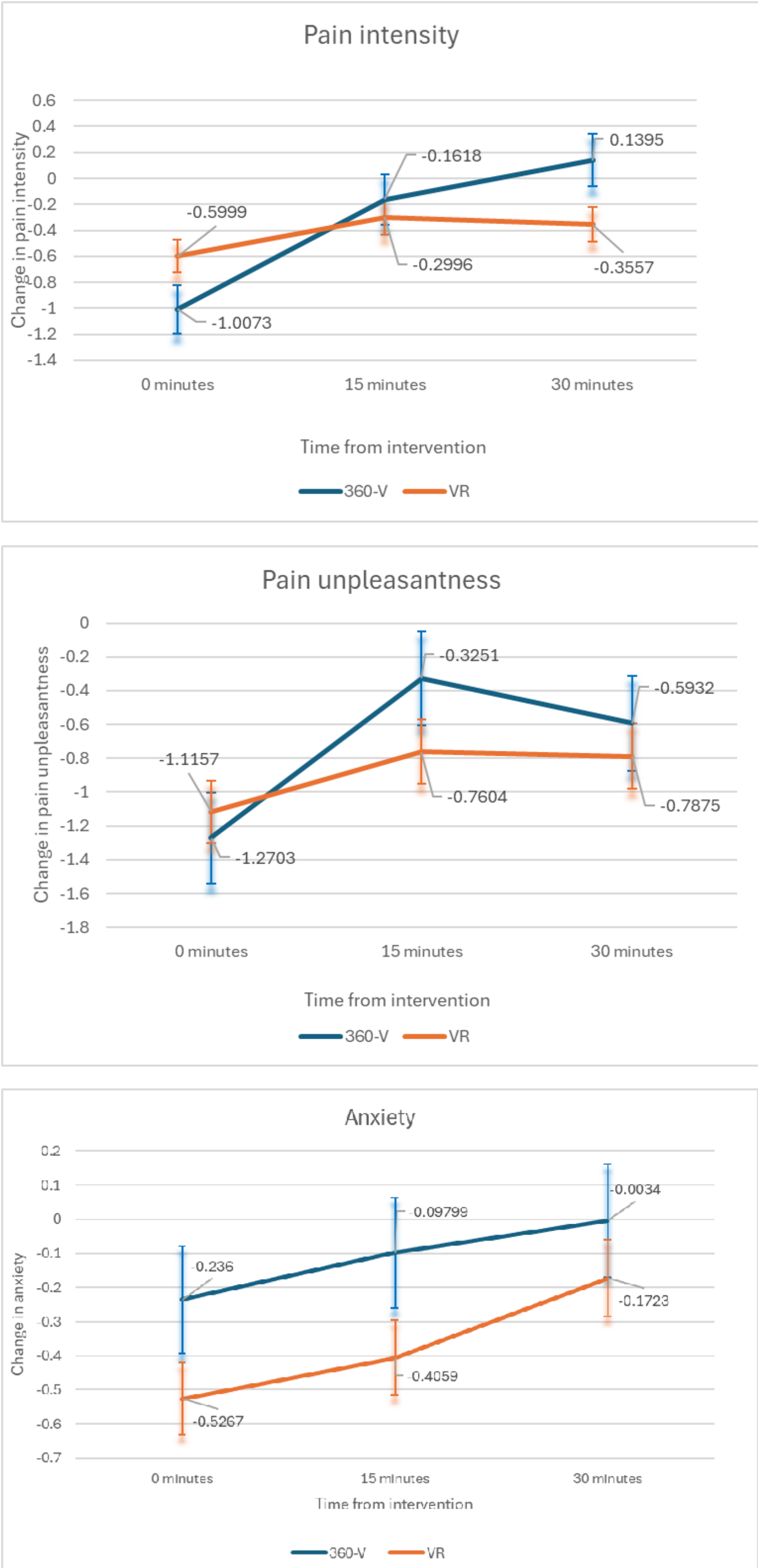
[c]CASI: Child Anxiety Sensitivity Index.

Changes From Baseline (VR vs Control)

Pain Intensity

Patients who participated in VR reported significantly decreased pain intensity from baseline (0.41 more decrease in pain intensity with SE 0.23) compared with those in the 360-V group at 30 minutes (*P*=.04) before multiplicity adjustment but not after multiplicity adjustment. There was no significant difference from baseline in reported pain intensity between VR vs 360-V immediately following the session (*P*=.08) or after 15 minutes (*P*=.56; Figure 3).

Figure 3. Changes in baseline in pain intensity, pain unpleasantness, and anxiety in time points following 360° video (360-V) and virtual reality (VR) in a mixed effect model with standard error bars.



Pain Unpleasantness

There was no significant difference in the reported pain unpleasantness between patients who participated in VR versus 360-V immediately following the VR session (*P*=.64), after 15 minutes (*P*=.20), or after 30 minutes (*P*=.57; Figure 3).

Anxiety

There were no significant differences in reported anxiety from baseline between patients who participated in the VR versus 360-V immediately following the session (*P*=.13), after 15 minutes (*P*=.12), or after 30 minutes (*P*=.40; Figure 3).

Inpatient Pain and Opioid Use

There were no significant differences in mean AUC pain scores between VR and 360-V (*P*=.60). There were also no significant differences in inpatient opioid use (MME/kg/day) between VR and 360-V (*P*=.26; Table 3).

Table . Inpatient pain and opioid use.

Characteristic	VR[a], mean (SD)	360-V[b], mean (SD)	*P* value[c]
Inpatient pain (AUC[d])			.60
POD[e]1	110.0 (34.3)	107.9 (31.8)	
POD2	95.9 (32.5)	90.6 (36.0)	
Inpatient opioid use (MME[f]/kg/day)			.26
POD0	0.18 (0.29)	0.18 (0.24)	
POD1	0.58 (0.22)	0.64 (0.28)	
POD2	0.50 (0.17)	0.51 (0.20)	

[a]VR: virtual reality.
[b]360-V: 360° video.
[c]*P* value from mixed effects models.
[d]AUC: area under the curve.
[e]POD: postoperative day.
[f]MME: morphine milligram equivalents.

Discussion

Principal Results

In our study, we found that active, immersive VR experiences had some trends to transient effects on both acute pain and anxiety compared to a nonimmersive 360-V control; however, these effects did not meet statistical significance. Patients who participated in VR reported a significantly greater decrease in pain intensity from baseline (0.41 with SE 0.23) compared with those in the 360-V group at 30 minutes (*P*=.04) before multiplicity adjustment but not after multiplicity adjustment. The trends in reduction in pain and anxiety were small; these trends did not achieve clinical significance either. Current literature indicates that a reduction of at least 2 points on the NRS for pain intensity or a 30% reduction in pain is considered clinically significant [39]. We did not see effects on overall AUC pain scores or opioid use. In this research, the VR-GR experience was likely distraction-based, as we could not document or assess feedback on patients' ability to perform the guided relaxation techniques correctly. In spite of not reaching statistical significance, these trends are not an absence of evidence of the effectiveness of VR to reduce postoperative pain and anxiety. We had relatively similar treatment conditions in small samples with attrition. The trends point us in the direction of future work.

Demographically, our two groups showed no significant differences except for a difference in PCS scores (VR: median 18, IQR 15 - 23; VR-360: median 24, IQR 16 - 28; *P*=.04). However, this result, while statistically significant, may not be clinically significant, as only PCS scores above 30 are clinically relevant, and neither group's median score exceeded 30 [40].

Although we noted some trends toward reduction in acute pain and anxiety from immersive VR following Nuss repair of pectus excavatum, these effects did not result in a significant change in AUC pain scores or inpatient opioid usage between the two groups. Several factors may account for the lack of significance. At our institution, the standard postoperative pain management protocol for following pectus surgery involves the scheduled administration of opioids, meaning all patients receive a standardized, weight-based dosage of opioids during their hospital stay, regardless of their actual pain level, with variations only in as-needed doses [10]. Consequently, opioid consumption may not accurately represent the patients' pain levels and opioid requirements. Home medication use might better reflect patients' pain and opioid needs. This study limited VR experiences to hospitalized patients. Extending its use past hospital discharge may have yielded different results. Future studies should investigate the integration of VR therapy into postoperative pain management both during and after hospitalization.

Limitations

Although this study was a prospective, randomized clinical trial, which can provide the best clinical evidence and support for VR, it has several limitations due to both study design and factors outside the control of the research team. Our patient population was somewhat homogeneous, as most individuals undergoing pectus excavatum repair are adolescent white males [38]. Therefore, our findings may not be generalizable to a broader population. Additionally, our control group may have been too similar to our intervention group. While the 360-V group did not receive audio instructions for guided relaxation, they still used a VR headset and experienced some level of distraction and immersion. Hence, the difference between the two groups may not have been substantial enough to detect a meaningful difference. This lack of significant differences in the treatment groups may account for the lack of statistical differences between the two groups when comparing pain and anxiety. Future research will use nonimmersive, non-headset control and retrospectively compare historical data to better assess the impact of VR on these outcomes.

Reporting bias may have also played a role, as study participants may have felt inclined to report decreased pain and/or anxiety after treatment due to their perception of receiving an intervention, regardless of actual changes. The self-reported 1 to 10 rating scale is limited and has been shown to have a moderate correlation with clinical indicators of pain; thus, we cannot rely exclusively on this measure for pain evaluation [41]. Future studies should consider additional outcome measures.

The limited use of the intervention (10 min per d) makes it unlikely to produce meaningful changes in the severe postoperative pain often experienced by these patients. This suggests a need for further investigation of VR for pain relief, potentially incorporating VR interventions more systematically throughout the perioperative period. This could include preoperative VR exposure, repeated interventions multiple times per day, for additional consecutive days following surgery. We accounted for our attrition in our statistical analysis. However, the attrition with small groups to start likely was one reason for our inability to find statistically significant differences in results between the two groups.

Comparison With Prior Work

Controlling pain after surgery is important; uncontrolled postoperative pain can lead to increased morbidity, decreased function, prolonged recovery, and higher costs [42]. Severe acute postoperative pain can also lead to chronic postoperative pain, with rates of chronic postsurgical pain reported to be about 20% in pediatric populations [43]. Furthermore, opioid use after surgery also has risks, including persistent opioid use postoperatively; one study found a rate of 4.8% of persistent opioid use in postoperative adolescents as compared to a 0.1% rate of persistent opioid use in their nonsurgical matched cohorts [44].

Our study's results align with prior research suggesting that VR-D may play a role in transiently reducing acute pain and that relaxation via guided imagery can promote reductions in both pain and anxiety [20,45]. Research consistently shows that VR is effective in lowering procedural pain, anxiety, and fear in pediatric patients, particularly during needle-related and other painful interventions [46-55]. A few studies have also demonstrated the feasibility of using VR to alleviate acute postoperative pain [21,56]. In adults, a meta-analysis found that patients receiving perioperative VR had lower pain scores than those receiving usual care (mean deviation −0.64, 95% CI −1.05 to −0.22; P<.02). Additionally, patients receiving VR postoperatively experienced a significant reduction in pain scores (mean deviation −0.50, 95% CI −0.76 to −0.24; P=.002) [57]. One pediatric study indicated that a single preoperative VR experience reduced the need for rescue analgesics in the recovery unit for painful procedures [58].

Relaxation-guided imagery has been shown to reduce both pain and anxiety in children undergoing minor surgery [59]. VR-GR has also demonstrated effectiveness in reducing pain and anxiety in children during medical procedures. These effects were immediate but transient, with some studies reporting reductions lasting up to 30 minutes after a session [21,60,61]. Additionally, a small (n=51) single-center, prospective study evaluated a single VR-GR session for acute postoperative pain and anxiety in children and adolescents. This study showed similar transient reductions in pain intensity and anxiety [21].

Using guided relaxation, our study aimed to harness the benefits of using mind-body techniques to manage postoperative pain. However, there are several potential reasons for the lack of clinically significant lasting effects on acute pain and anxiety in this population or opioid consumption. First, our study did not include a non-headset control group. Previous studies assessing the effects of VR on opioid use compared VR experiences to standard care without VR [55,62,63]. In our study, every patient utilized a VR headset and had some level of an immersive experience, even with 360-V (a nonimmersive option). As a result, all patients, including the control group, likely experienced some distraction.

Moreover, the effectiveness of the VR sessions may have been limited by the short duration of each session and the low number of sessions. Ten minutes per day may be an inadequate time to achieve lasting pain relief of severe acute pain, especially as compared to other therapeutic and/or pharmacologic interventions that are administered more frequently. Future studies should investigate the optimal timing and length of intervention for maximal benefit of VR.

While our results indicated that active, immersive VR experiences had some trends toward effects on both acute pain and anxiety compared to the nonimmersive 360-V control, these effects were not clinically significant. The study may not have produced clear positive results due to experimental factors such as the similarity of the VR and control treatments, the short duration of treatments, and the protocol-driven high use of opioids that were not adequately considered in the study design. Nevertheless, the findings provide a valuable framework for designing future VR studies. We can learn from our null results to design future VR studies with a control treatment that does not use a VR headset and use a population that only receives opioids on an as-needed basis. Longer treatment duration and less subject attrition could also lead to more significant results.

Conclusions

This study found that daily, 10-minute VR sessions had trends toward transiently reducing pain and anxiety compared to a 360-V experience in participants following Nuss repair of pectus excavatum. These results were not clinically significant. Due to the limited duration of the intervention and the standardized, scheduled, high utilization of opioids in this population, VR was not sufficient in significantly decreasing opioid use and overall AUC pain scores. Despite these conclusions, exploring guided relaxation VR as an adjunct to, rather than a replacement for, postoperative pharmacologic analgesics may prove valuable. Increasing the length and frequency of VR experiences per day, along with a policy of not automatically administering opioids unless requested, may help decrease opioid usage and AUC pain scores. A systematic integration of VR into perioperative care is likely necessary to impact the pain trajectory and opioid usage in postoperative patients. Furthermore, improving the VR experience to incorporate true guided relaxation would likely enhance effectiveness compared to a purely distraction-based approach. Future studies are needed to further explore the use of this therapy in postoperative pain management.

Acknowledgments

The authors gratefully acknowledge Maria E Ashton, MS, RPh, MBA, medical writer, Department of Anesthesiology, Cincinnati Children's Hospital Medical Center, for providing writing assistance, editing, and proofreading.

The authors gratefully acknowledge Luke Wilson and the Stanford CHARIOT program for the development and use of the Mindful Aurora application.

Funding

Support was provided by the Department of Anesthesiology at Cincinnati Children's Hospital Medical Center and a Research Innovation/Pilot Award from Cincinnati Children's Hospital Medical Center.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1122 KB - periop_v9i1e80902_app1.pdf]

References

1. Yaster M, McNaull PP, Davis PJ. The opioid epidemic in pediatrics: a 2020 update. Curr Opin Anaesthesiol 2020 Jun;33(3):327-334. [doi: 10.1097/ACO.0000000000000865] [Medline: 32371640]
2. Drug overdose deaths: facts and figures. National Institute on Drug Abuse (NIDA). URL: https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates [accessed 2026-02-16]
3. Hudgins JD, Porter JJ, Monuteaux MC, Bourgeois FT. Trends in opioid prescribing for adolescents and young adults in ambulatory care settings. Pediatrics 2019 Jun;143(6):e20181578. [doi: 10.1542/peds.2018-1578] [Medline: 31138669]
4. Johnston LD, O'Malley PM, et al. Monitoring the future" national survey results on drug use, 1975–2017: 2017 overview, key findings on adolescent drug use. : Institute for Social Research, University of Michigan; 2017 URL: https://monitoringthefuture.org/wp-content/uploads/2022/08/mtf-overview2017.pdf [accessed 2026-02-16]
5. Groenewald CB, Zhou C, Palermo TM, Van Cleve WC. Associations between opioid prescribing patterns and overdose among privately insured adolescents. Pediatrics 2019 Nov;144(5):e20184070. [doi: 10.1542/peds.2018-4070] [Medline: 31575622]
6. Battaglia M, Groenewald CB, Campbell F, et al. We need to talk: the urgent conversation on chronic pain, mental health, prescribing patterns and the opioid crisis. J Psychopharmacol 2023 May;37(5):437-448. [doi: 10.1177/02698811221144635] [Medline: 37171242]
7. Breglio AM, Fitzgerald TN, Moore CB, Einhorn LM. Evaluation of analgesic practice changes following the Nuss procedure in pediatric patients. J Surg Res 2023 Nov;291:289-295. [doi: 10.1016/j.jss.2023.06.018] [Medline: 37481964]
8. Densmore JC, Peterson DB, Stahovic LL, et al. Initial surgical and pain management outcomes after Nuss procedure. J Pediatr Surg 2010 Sep;45(9):1767-1771. [doi: 10.1016/j.jpedsurg.2010.01.028] [Medline: 20850618]
9. Lai K, Notrica DM, McMahon LE, et al. Cryoablation in 350 Nuss procedures: evolution of hospital length of stay and opioid use. J Pediatr Surg 2023 Aug;58(8):1435-1439. [doi: 10.1016/j.jpedsurg.2022.10.051] [Medline: 36494205]
10. Walter CM, Lee CS, Moore DL, et al. Retrospective study comparing outcomes of multimodal epidural and erector spinae catheter pain protocols after pectus surgery. J Pediatr Surg 2023 Mar;58(3):397-404. [doi: 10.1016/j.jpedsurg.2022.06.017] [Medline: 35907711]
11. Tashjian VC, Mosadeghi S, Howard AR, et al. Virtual reality for management of pain in hospitalized patients: results of a controlled trial. JMIR Ment Health 2017 Mar 29;4(1):e9. [doi: 10.2196/mental.7387] [Medline: 28356241]
12. Simonetti V, Tomietto M, Comparcini D, Vankova N, Marcelli S, Cicolini G. Effectiveness of virtual reality in the management of paediatric anxiety during the peri-operative period: a systematic review and meta-analysis. Int J Nurs Stud 2022 Jan;125:104115. [doi: 10.1016/j.ijnurstu.2021.104115] [Medline: 34781118]
13. Olbrecht VA, O'Conor KT, Williams SE, et al. Transient reductions in postoperative pain and anxiety with the use of virtual reality in children. Pain Med 2021 Nov 26;22(11):2426-2435. [doi: 10.1093/pm/pnab209] [Medline: 34175959]

14. Melzack R, Wall PD. Pain mechanisms: a new theory. Science 1965 Nov 19;150(3699):971-979. [doi: 10.1126/science.150.3699.971] [Medline: 5320816]
15. Smith KL, Wang Y, Colloca L. Impact of virtual reality technology on pain and anxiety in pediatric burn patients: a systematic review and meta-analysis. Front Virtual Real 2022 Jan;2:2022. [doi: 10.3389/frvir.2021.751735] [Medline: 36247202]
16. Indovina P, Barone D, Gallo L, Chirico A, De Pietro G, Giordano A. Virtual reality as a distraction intervention to relieve pain and distress during medical procedures: a comprehensive literature review. Clin J Pain 2018 Sep;34(9):858-877. [doi: 10.1097/AJP.0000000000000599] [Medline: 29485536]
17. Ahmadpour N, Randall H, Choksi H, Gao A, Vaughan C, Poronnik P. Virtual reality interventions for acute and chronic pain management. Int J Biochem Cell Biol 2019 Sep;114:105568. [doi: 10.1016/j.biocel.2019.105568] [Medline: 31306747]
18. Hadjiat Y, Marchand S. Virtual reality and the mediation of acute and chronic pain in adult and pediatric populations: research developments. Front Pain Res (Lausanne) 2022;3:840921. [doi: 10.3389/fpain.2022.840921] [Medline: 35599969]
19. Sánchez-Caballero E, Ortega-Donaire L, Sanz-Martos S. Immersive virtual reality for pain and anxiety management associated with medical procedures in children and adolescents: a systematic review. Children (Basel) 2024 Aug 13;11(8):975. [doi: 10.3390/children11080975] [Medline: 39201910]
20. Hoffman HG, Rodriguez RA, Gonzalez M, et al. Immersive virtual reality as an adjunctive non-opioid analgesic for pre-dominantly Latin American children with large severe burn wounds during burn wound cleaning in the intensive care unit: a pilot study. Front Hum Neurosci 2019;13(262):262. [doi: 10.3389/fnhum.2019.00262] [Medline: 31440148]
21. Olbrecht VA, O'Conor KT, Williams SE, et al. Guided relaxation–based virtual reality for acute postoperative pain and anxiety in a pediatric population: pilot observational study. J Med Internet Res 2021 Jul 12;23(7):e26328. [doi: 10.2196/26328] [Medline: 34048358]
22. Hemington KS, Cheng JC, Bosma RL, Rogachov A, Kim JA, Davis KD. Beyond negative pain-related psychological factors: resilience is related to lower pain affect in healthy adults. J Pain 2017 Sep;18(9):1117-1128. [doi: 10.1016/j.jpain.2017.04.009] [Medline: 28506777]
23. Agoston AM, Sieberg CB. Nonpharmacologic treatment of pain. Semin Pediatr Neurol 2016 Aug;23(3):220-223. [doi: 10.1016/j.spen.2016.10.005] [Medline: 27989329]
24. Turgut A, Özcan İlçe A, Öztürk H. The effect of immersive virtual reality application on anxiety, pain, and parental satisfaction in the perioperative process of children: a randomized controlled trial. Pain Manag Nurs 2024 Dec;25(6):584-590. [doi: 10.1016/j.pmn.2024.06.002] [Medline: 38955550]
25. Orgil Z, Johnson L, Karthic A, et al. Feasibility and acceptability of perioperative application of biofeedback-based virtual reality versus active control for pain and anxiety in children and adolescents undergoing surgery: protocol for a pilot randomised controlled trial. BMJ Open 2023 Jan 25;13(1):e071274. [doi: 10.1136/bmjopen-2022-071274] [Medline: 36697053]
26. Olbrecht VA, Williams SE, O'Conor KT, et al. Guided relaxation-based virtual reality versus distraction-based virtual reality or passive control for postoperative pain management in children and adolescents undergoing Nuss repair of pectus excavatum: protocol for a prospective, randomised, controlled trial (FOREVR Peds trial). BMJ Open 2020 Dec 30;10(12):e040295. [doi: 10.1136/bmjopen-2020-040295] [Medline: 33380482]
27. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
28. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomized trials. JAMA 2025 Jun 10;333(22):1998-2005. [doi: 10.1001/jama.2025.4347] [Medline: 40228499]
29. Tsze DS, von Baeyer CL, Pahalyants V, Dayan PS. Validity and reliability of the verbal numerical rating scale for children aged 4 to 17 years with acute pain. Ann Emerg Med 2018 Jun;71(6):691-702. [doi: 10.1016/j.annemergmed.2017.09.009] [Medline: 29107409]
30. Price TJ, Ray PR. Recent advances toward understanding the mysteries of the acute to chronic pain transition. Curr Opin Physiol 2019 Oct;11:42-50. [doi: 10.1016/j.cophys.2019.05.015] [Medline: 32322780]
31. Pielech M, Ryan M, Logan D, Kaczynski K, White MT, Simons LE. Pain catastrophizing in children with chronic pain and their parents: proposed clinical reference points and reexamination of the Pain Catastrophizing Scale measure. Pain 2014 Nov;155(11):2360-2367. [doi: 10.1016/j.pain.2014.08.035] [Medline: 25180013]
32. Silverman WK, Fleisig W, Rabian B, Peterson RA. Childhood anxiety sensitivity index. J Clin Child Psychol 1991 Jun;20(2):162-168. [doi: 10.1207/s15374424jccp2002_7]
33. Gold JI, Mahrer NE. Is virtual reality ready for prime time in the medical space? A randomized control trial of pediatric virtual reality for acute procedural pain management. J Pediatr Psychol 2018 Apr 1;43(3):266-275. [doi: 10.1093/jpepsy/jsx129] [Medline: 29053848]
34. Proczkowska M, Ericsson E. Validity of the modified-distraction-short-scale and verbal-numeric-anxiety-fear-rating-scale for children in a preoperative setting. Paediatr Anaesth 2024 Feb;34(2):121-129. [doi: 10.1111/pan.14765] [Medline: 37728169]

35. Adams MCB, Sward KA, Perkins ML, Hurley RW. Standardizing research methods for opioid dose comparison: the NIH HEAL morphine milligram equivalent calculator. Pain 2025 Feb 3;166(8):1729-1737. [doi: 10.1097/j.pain.0000000000003529] [Medline: 39907478]
36. Morphine milligram equivalent (MME) calculator: frequently asked questions (FAQs) for HEAL researchers. : National Institutes of Health; 2025 URL: https://www.nih.gov/sites/default/files/2026-01/MME%20Calculator%20FAQ_508.pdf [accessed 2026-02-16]
37. Pitt JB, Zeineddin S, Carter M, et al. Demographics of anterior chest wall deformity patients: a tertiary children's hospital experience. J Surg Res 2024 Jan;293:451-457. [doi: 10.1016/j.jss.2023.09.040] [Medline: 37827024]
38. Mack SJ, Till BM, Huang C, et al. National trends in pectus excavatum repair: patient age, facility volume, and outcomes. J Thorac Dis 2022 Apr;14(4):952-961. [doi: 10.21037/jtd-21-1671] [Medline: 35572899]
39. Rowbotham MC. What is a "clinically meaningful" reduction in pain? Pain 2001 Nov;94(2):131-132. [doi: 10.1016/S0304-3959(01)00371-2] [Medline: 11690725]
40. Sullivan MJ. The pain catastrophizing scale: user manual. : McGill University; 2009 URL: https://drsandler.in/PCSmanual.pdf [accessed 2026-02-16]
41. Nugent SM, Lovejoy TI, Shull S, Dobscha SK, Morasco BJ. Associations of pain numeric rating scale scores collected during usual care with research administered patient reported pain outcomes. Pain Med 2021 Oct 8;22(10):2235-2241. [doi: 10.1093/pm/pnab110] [Medline: 33749760]
42. Gan TJ. Poorly controlled postoperative pain: prevalence, consequences, and prevention. J Pain Res 2017;10:2287-2298. [doi: 10.2147/JPR.S144066] [Medline: 29026331]
43. Rabbitts JA, Fisher E, Rosenbloom BN, Palermo TM. Prevalence and predictors of chronic postsurgical pain in children: a systematic review and meta-analysis. J Pain 2017 Jun;18(6):605-614. [doi: 10.1016/j.jpain.2017.03.007] [Medline: 28363861]
44. Harbaugh CM, Lee JS, Hu HM, et al. Persistent opioid use among pediatric patients after surgery. Pediatrics 2018 Jan;141(1):e20172439. [doi: 10.1542/peds.2017-2439] [Medline: 29203521]
45. Chen SF, Wang HH, Yang HY, Chung UL. Effect of relaxation with guided imagery on the physical and psychological symptoms of breast cancer patients undergoing chemotherapy. Iran Red Crescent Med J 2015 Nov;17(11):e31277. [doi: 10.5812/ircmj.31277] [Medline: 26734485]
46. Dumoulin S, Bouchard S, Ellis J, et al. A randomized controlled trial on the use of virtual reality for needle-related procedures in children and adolescents in the emergency department. Games Health J 2019 Aug;8(4):285-293. [doi: 10.1089/g4h.2018.0111] [Medline: 31135178]
47. Iannicelli AM, Vito D, Dodaro CA, et al. Does virtual reality reduce pain in pediatric patients? A systematic review. Ital J Pediatr 2019 Dec 30;45(1):171. [doi: 10.1186/s13052-019-0757-0] [Medline: 31888710]
48. Eijlers R, Utens EMWJ, Staals LM, et al. Systematic review and meta-analysis of virtual reality in pediatrics: effects on pain and anxiety. Anesth Analg 2019 Nov;129(5):1344-1353. [doi: 10.1213/ANE.0000000000004165] [Medline: 31136330]
49. Czech O, Wrzeciono A, Rutkowska A, Guzik A, Kiper P, Rutkowski S. Virtual reality interventions for needle-related procedural pain, fear and anxiety-a systematic review and meta-analysis. J Clin Med 2021 Jul 23;10(15):3248. [doi: 10.3390/jcm10153248] [Medline: 34362032]
50. Walther-Larsen S, Petersen T, Friis SM, Aagaard G, Drivenes B, Opstrup P. Immersive virtual reality for pediatric procedural pain: a randomized clinical trial. Hosp Pediatr 2019 Jul;9(7):501-507. [doi: 10.1542/hpeds.2018-0249] [Medline: 31160472]
51. Jenabi E, Bashirian S, Salehi AM, Rafiee M, Bashirian M. Virtual reality for pain reduction during intravenous injection in pediatrics: a systematic review and meta-analysis of controlled clinical trials. Clin Exp Pediatr 2023 Dec;66(12):533-537. [doi: 10.3345/cep.2022.01193] [Medline: 37321586]
52. Addab S, Hamdy R, Thorstad K, Le May S, Tsimicalis A. Use of virtual reality in managing paediatric procedural pain and anxiety: an integrative literature review. J Clin Nurs 2022 Nov;31(21-22):3032-3059. [doi: 10.1111/jocn.16217] [Medline: 35068011]
53. Alshatrat SM, Sabarini JM, Hammouri HM, Al-Bakri IA, Al-Omari WM. Effect of immersive virtual reality on pain in different dental procedures in children: a pilot study. Int J Paediatr Dent 2022 Mar;32(2):264-272. [doi: 10.1111/ipd.12851] [Medline: 34133809]
54. Gao Y, Xu Y, Liu N, Fan L. Effectiveness of virtual reality intervention on reducing the pain, anxiety and fear of needle-related procedures in paediatric patients: a systematic review and meta-analysis. J Adv Nurs 2023 Jan;79(1):15-30. [doi: 10.1111/jan.15473] [Medline: 36330583]
55. Camacho-Cruz J, Palacios-Ariza MA, Orrego-Celestino L, et al. Effectiveness of non-immersive virtual reality in the management of procedure-related pain in preschool children: a randomized clinical trial. Eur J Pediatr 2023 Sep;182(9):4103-4112. [doi: 10.1007/s00431-023-05070-5] [Medline: 37407711]
56. Mosso JL, Rizzo S, Wiederhold B, et al. Cybertherapy--new applications for discomfort reductions. Surgical care unit of heart, neonatology care unit, transplant kidney care unit, delivery room-cesarean surgery and ambulatory surgery, 27 case reports. Stud Health Technol Inform 2007;125:334-336. [Medline: 17377297]
57. Ding L, Hua H, Zhu H, et al. Effects of virtual reality on relieving postoperative pain in surgical patients: a systematic review and meta-analysis. Int J Surg 2020 Oct;82:87-94. [doi: 10.1016/j.ijsu.2020.08.033] [Medline: 32882400]

58. Eijlers R, Dierckx B, Staals LM, et al. Virtual reality exposure before elective day care surgery to reduce anxiety and pain in children: a randomised controlled trial. Eur J Anaesthesiol 2019 Oct;36(10):728-737. [doi: 10.1097/EJA.0000000000001059] [Medline: 31356373]
59. Vagnoli L, Bettini A, Amore E, De Masi S, Messeri A. Relaxation-guided imagery reduces perioperative anxiety and pain in children: a randomized study. Eur J Pediatr 2019 Jun;178(6):913-921. [doi: 10.1007/s00431-019-03376-x] [Medline: 30944985]
60. Bernaerts S, Bonroy B, Daems J, et al. Virtual reality for distraction and relaxation in a pediatric hospital setting: an interventional study with a mixed-methods design. Front Digit Health 2022;4:866119. [doi: 10.3389/fdgth.2022.866119] [Medline: 35712230]
61. Hoag JA, Karst JS, Bingen KM, Palou-Torres A, Yan K. Distracting through procedural pain and distress using virtual reality and guided imagery in pediatric, adolescent, and young adult patients: randomized controlled trial. J Med Internet Res 2022 Apr 18;24(4):e30260. [doi: 10.2196/30260] [Medline: 35436209]
62. Xiang H, Shen J, Wheeler KK, et al. Efficacy of smartphone active and passive virtual reality distraction vs standard care on burn pain among pediatric patients: a randomized clinical trial. JAMA Netw Open 2021 Jun 1;4(6):e2112082. [doi: 10.1001/jamanetworkopen.2021.12082] [Medline: 34152420]
63. Gold JI, SooHoo M, Laikin AM, Lane AS, Klein MJ. Effect of an immersive virtual reality intervention on pain and anxiety associated with peripheral intravenous catheter placement in the pediatric setting: a randomized clinical trial. JAMA Netw Open 2021 Aug 2;4(8):e2122569. [doi: 10.1001/jamanetworkopen.2021.22569] [Medline: 34432011]

Abbreviations

360-V: 360° video
AUC: area under the curve
CCHMC : Cincinnati Children’s Hospital Medical Center
CONSORT: Consolidated Standard of Reporting Trials
MME: morphine milligram equivalents
NRS: numerical rating scale
PCS : Pain Catastrophizing Scale
POD : postoperative day
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
VR: virtual reality
VR-D : distraction-based gaming virtual reality
VR-DGR: distraction-based guided relaxation virtual reality
VR-GR: guided relaxation–based virtual reality

Edited by N Rohatgi; submitted 18.Jul.2025; peer-reviewed by I Cornelisz, JW Gander; revised version received 20.Jan.2026; accepted 06.Feb.2026; published 10.Mar.2026.

Please cite as:
Walter CM, Froass D, Bell N, Haack L, Boehmer C, Bruguera Torres C, Spivak R, Chou M, Geisler K, O'Conor K, Williams SE, Ding L, King CD, Olbrecht VA
Virtual Reality for the Management of Postoperative Pain and Anxiety in Children and Adolescents Undergoing Nuss Repair of Pectus Excavatum: Randomized Controlled Trial
JMIR Perioper Med 2026;9:e80902
URL: https://periop.jmir.org/2026/1/e80902
doi:10.2196/80902

A Novel Customizable Datamart and Tableau Dashboard to Monitor Multiple Enhanced Recovery After Surgery Programs: Development and Validation Study

Sunitha Margaret Singh[1], MD; Susannah Oster[2], BA; Efrat Bolze[3], BA; Aaron Sasson[4], MD; James Nicholson[5], MD; Elliott Bennett-Guerrero[2], MD

[1]Department of Perioperative Surgical Services, Stony Brook University Medical Center, 101 Nichols Road, Stony Brook, NY, United States
[2]Department of Anesthesiology, Stony Brook University Medical Center, Stony Brook, NY, United States
[3]Enterprise Analytics, Stony Brook Medicine Information Technology (SBMIT), St. James, NY, United States
[4]Department of Surgery, Stony Brook University Medical Center, Stony Brook, NY, United States
[5]Department of Orthopedics, Stony Brook University Medical Center, Stony Brook, NY, United States

Corresponding Author:
Sunitha Margaret Singh, MD
Department of Perioperative Surgical Services, Stony Brook University Medical Center, 101 Nichols Road, Stony Brook, NY, United States

Abstract

Background: Enhanced recovery after surgery (ERAS) programs bundle evidence-based interventions to standardize care, expedite recovery, and improve outcomes. As ERAS programs have expanded, it has become clear that a major challenge is monitoring the compliance of bundle elements and outcomes to feedback performance to stakeholders and guide changes. Manual data abstraction is onerous and not feasible. Reliance on receiving new reports from busy health system IT groups is challenging. Therefore, we sought to address this unmet need at our hospital by developing a novel ERAS Datamart system.

Objective: Our objectives were to develop a novel Datamart and Tableau dashboard to (1) enable continuous analysis of data, harvested directly from the electronic medical record (EMR), measure compliance and outcomes, and (2) enable end users (e.g., an ERAS coordinator) to create reports customized based on surgical procedure types, requested data variables, and custom date ranges.

Methods: After “buy-in” from hospital leadership and other stakeholders, data metrics were identified and categorized according to phase of care, that is, preoperative, intraoperative, and postoperative. A multidisciplinary team reviewed *International Classification of Diseases, Tenth Revision* procedure codes to capture EMR data for patients undergoing ERAS procedures. IT was given a master list with metric names, definitions, and screenshots of the discrete field in the EMR to assist with building the metrics. Validations of the novel Datamart were done against known ERAS patient populations maintained by the surgery clinic.

Results: The Datamart and Tableau dashboard has been built, is functional, and contains over 17,000 patients across 5 ERAS service lines: colorectal (n=1742), joint replacement (n=4235), surgical oncology (n=941), bariatric (n=1130), and cesarean section (n=9390). Currently, 56 metrics spanning the perioperative period have been validated across these populations. Reports can be tailored according to patients, time frames, and metrics. If desired, patient-level raw data can be exported for statistical analyses. Two use cases (total joint replacement and surgical oncology ERAS programs) are presented showing how the Datamart can be used.

Conclusions: Discrete fields within an EMR can be successfully captured into a novel Datamart and visualized using a custom Tableau dashboard for providing stakeholder feedback, facilitating quality improvement analyses, and auditing pathways.

(JMIR Perioper Med 2026;9:e82472) doi:10.2196/82472

KEYWORDS

data monitoring; enhanced recovery; enhanced recovery after surgery; ERAS; perioperative outcomes; quality improvement; web platform

Introduction

Enhanced recovery after surgery (ERAS) programs have transformed perioperative care by implementing evidence-based interventions that aim to standardize patient care and management, decrease resource utilization, expedite recovery, and improve patient outcomes [1-3]. The success and efficacy of ERAS programs are most likely achieved through the implementation of a comprehensive approach that bundles care for patients undergoing elective surgery, encompassing approximately 20 care elements [1,4].

It is increasingly recognized that as ERAS programs increase in size, it is very challenging to monitor and track bundle elements to feedback performance and guide outcomes. Traditional methods usually rely on manual data abstraction, frequent reports generated by hospital IT systems, or the use of third-party data warehouses. As ERAS programs grow in size and complexity, manual data abstraction becomes impractical due to time demands, error risk, and challenges with real-time analysis. Basic data points, such as length of stay, are easier to track, but capturing complex metrics, such as total opioid use (e.g., oral morphine equivalents), is often not feasible. Reliance on an IT report strategy is typically limited by very long delays in obtaining data reports from hospital IT workers who are usually burdened with many requests. The use of third-party data warehouses, for example, ERAS Interactive Audit System (EIAS), offers an alternative but raises concerns about data security, control, costs, system downtimes, and limited flexibility [5]. To address the above limitations, our institution created a novel dynamic Datamart dashboard.

Methods

Overview

Stony Brook University Hospital is a tertiary care academic medical center on Long Island, New York. Its first ERAS program (lumbar spine fusion) began in 2016, and an additional 9 ERAS programs were subsequently added. As the program grew, this revealed the unmet need for how to efficiently capture and monitor compliance and outcomes across a large number of patients.

As described in more detail below, the process for creating this system included: (1) leadership support and data governance, (2) validated identification of relevant patients to be included in the Datamart, (3) metric identification and validation, and (4) Tableau visualization as the user interface.

Leadership Support and Data Governance

Under an institutional quality assurance program, a guideline was developed to map the creation of a novel Datamart and Tableau dashboard and govern the data extracted. To prioritize this effort, a value statement was presented to institutional leadership. This statement provided background information on the institution's ERAS programs, highlighted their prior success, and outlined the intended purpose of the Datamart and Tableau dashboard, such as monitoring and improving compliance, reducing errors associated with manual data abstraction, and limiting frequent report requests made to IT.

After approval, a Global ERAS Data Governance plan was conceived with policies and procedures for protecting and using ERAS data. These included how the data would be stored and protected, who would have access to data, and how data would be managed (e.g., requests for aggregate and patient-level reports, quality assurance (QA) analyses, and institutional review board–approved research projects). The Global ERAS Data Governance plan was subsequently signed by applicable departmental chairpersons to ensure data analysis was conducted in accordance with institutional standards and to protect against breaches of protected health information.

Identification of Relevant Patients

Choice of International Classification of Diseases, Tenth Revision Codes Methods

We considered several possible strategies for identifying relevant patients for a given ERAS pathway. The hospital's operating room schedule, that is, planned surgical procedure, provides information on the "planned" procedure; however, it does not accurately reflect the "actual" surgery performed. Current procedural terminology professional billing codes were not used since the hospital's IT department did not have direct access to them. Therefore, we decided to use the *International Classification of Diseases, Tenth Revision* (*ICD-10*) Procedure Coding System (PCS) since this was feasible and these codes are believed to be accurate as they are used for hospital billing purposes.

To ensure the accurate identification of the ERAS patient population, the team collaborated closely with surgical leads from each ERAS pathway. Surgical procedures were sent to the coding department to identify the *ICD-10* PCS associated with specific procedures. The coding department supplied the leading 4 digits of all *ICD-10* PCS for the specified procedures. These digits encompass the section, body part, root operation, and where relevant, the body part of the given procedure. This preliminary list was forwarded to the IT department, which then appended the remaining digits of the *ICD-10* PCS, corresponding to the approach, device, and qualifier, for the procedures performed. Subsequently, a multidisciplinary team reviewed the complete *ICD-10* PCS to ensure the accurate capture of electronic medical record (EMR) data for patients undergoing ERAS procedures. This task required careful attention due to the complexities and overlaps within *ICD-10* PCS across multiple procedure types.

To ensure a comprehensive approach, a strategy was developed that included both *ICD-10* procedure codes and additional criteria, such as surgical case procedure name and other associated details (eg, associated *ICD-10* PCS, ambulatory surgical center vs main operating room, surgeon). This established a robust, multistep process for accurately pinpointing the desired patient population. This method ensured that only patients who received care associated with an ERAS pathway were included, enhancing the precision and reliability of the data captured for monitoring and analysis.

Population Validation Methods

Next, we validated the accuracy of using these *ICD-10* procedure codes to identify the desired patient population. This validation aimed to identify instances where incorrect patients (non-ERAS patients) were erroneously included or, at the other extreme, ERAS patients were missing (i.e., not included). Admit type was then utilized to refine the patient population based on the urgency of admission, categorized as urgent, emergent, or elective. Since ERAS patients almost always fall into the “elective” category, this refinement allowed for an additional method for validating patient populations. Exception reports were generated and scheduled for automated reporting of patients who met the *ICD-10* PCS criteria but were admitted urgently or emergently. These records were then cross-referenced with the EMR to verify their eligibility for inclusion in the Datamart, ensuring only accurate ERAS populations were maintained. Several validations of the Datamart data were then performed by comparing the dataset against known ERAS patient populations maintained by the surgery clinic or stored in a REDCap (Research Electronic Data Capture) database.

Metric Identification and Validation Methods

ERAS programs usually include many (e.g., ≥20) best practice elements. For example, the use of nonopioid analgesics, such as acetaminophen and nonsteroidal medications, is prioritized to minimize opioid use [6]. Since there are no national benchmarks for these programs to identify metrics of interest, key stakeholders (e.g.surgeons, anesthesiologists, hospital quality department personnel) were engaged. Metrics were selected, defined, and aligned with institutional interest, key performance, and patient safety indicators, and the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) by these stakeholders. We also sought to identify metrics that would be broadly applicable to most elective surgical procedures since the proposed Datamart system would be used for all ERAS programs.

Given the variations in documentation that can occur across an institution, it was crucial to investigate how and when an EMR field was completed. For example, nursing staff across different surgical specialties might record specifics related to urinary catheterization in separate discrete fields within the EMR. Additionally, there could be discrepancies in other charting practices as well, such as documenting ambulation as the “number of steps” taken versus “number of laps” taken. Metrics were validated with each ERAS population to identify discrepancies and ensure that the discrete field identified could be applied to the majority of surgical populations. This validation of metrics was carried out in multiple phases (e.g., 6 - 10 metrics at a time with each population) to alleviate the burden of mass validation. Metric validations were performed against known ERAS populations using manual chart review to ensure accurate data extraction into the Datamart from discrete EMR fields.

Once finalized, the metrics were categorized according to phase of care: preoperative, intraoperative, postoperative, and discharge. IT was provided with metric names, definitions, and screenshots of the discrete fields in the EMR to assist with building the metrics. To compartmentalize the data in the Datamart, these phases were bucketed into 7 categories, encompassing various areas of care. These categories included patient characteristics, preoperative care, operating room, postoperative fluid, postoperative multimodal analgesia, postoperative opioid, and postoperative patient experience.

User Interface (Tableau Visualization) Methods

Tableau was chosen as the software for enabling a web-based dashboard user interface due to its on-demand filtering options and Health Insurance Portability and Accountability Act–compliant capabilities. An open query was established between our institution’s EMR relational database and the Tableau dashboard. This setup allowed automated data extracts from the EMR to the Tableau dashboard according to the desired export rate. Multiple viewpoints and designs were trialed through various builds created by IT. Essential filtering options, such as specified time frames, surgical specialty, surgical procedures, and other specific metrics, were established and selected based on their relevance and informativeness for patients receiving care related to ERAS.

Ethical Considerations

This project was conducted as part of an institutionally approved quality improvement/quality assurance initiative aimed at optimizing perioperative care processes and monitoring ERAS program performance. In accordance with our institution’s policies on human participants protections, quality improvement/quality assurance activities that are designed solely for internal program evaluation are not considered human participants research.

Results

Results are presented in the same sequence of events (1-4) as described in the Methods section.

Leadership Support and Data Governance

An ERAS data governance guidance plan was established. The document outlined the appropriate use of the ERAS data warehouse, including storage, protection from data breach and leak of protected health information, access to data, and ensuring data analysis is in accordance with institutional standards. Data access is managed by a data access group. Requests for data must be submitted in writing to the data access group for review and approval. Data can be used for quality and research scholarly activities.

Validation of Patient Selection Using *ICD-10* Procedures Codes

Initially, 50 metrics were trialed with *ICD-10* PCS for ERAS colorectal procedures. The patient population yielded from these codes returned many patients, more than 2-fold, compared to the known ERAS colorectal population. Preliminary validations revealed the need to (1) exclude surgery types such as emergent or urgent, (2) exclude overly broad or ambiguous *ICD 10* procedure codes, and (3) screen for potential inclusion of the planned surgical case procedure. After validating the colorectal population, the initial metrics were tested in the surgical oncology and joint replacement ERAS populations. We then backfilled data to 2015 to include pre-ERAS populations.

After final validations, the Datamart (last 6.5 y) contains more than 17,000 patients, consisting of colorectal (n=1742), surgical oncology (n=941), joint replacement (n=4235), bariatric (n=1130), and cesarian section (n=9390). In addition, the Datamart contains over 3000 archived cases from 2015 to 2018.

Metric Identification and Validation

Over 100 metrics of potential interest, spanning the perioperative period and including patient characteristics, were identified. Several of these metrics represent the same metric measured at different time points over several days, such as peak pain on postoperative day (POD) 0, peak pain on POD 1, and peak pain on POD 2. The metrics were categorized according to the phase of care: preoperative, intraoperative, and postoperative. Currently, there are 56 metrics (demographics, surgical and anesthesia care, and postoperative endpoints) in the Datamart Textbox 1.

Textbox 1. Current metrics active in the Datamart.

1. Patient characteristics
    - Patient age, mean or median
    - Body mass index, mean or median
    - Diabetes diagnosis, #/%
    - Current smoker, #/%
    - Chronic opioid use (regular opioid use as listed on home medications), #/%
    - Associated diagnosis code, cancer, #/%
    - Associated diagnosis code, inflammatory bowel disease, #/%
    - Actual procedure completed, #/%
2. Preoperative care
    - Preprocedural bowel prep, #/%
    - Preprocedural oral antibiotics, #/%
    - Total functional status score (sum of activities of daily living score on the day of surgery in the preoperative area), mean or median
    - Hemoglobin A_{1c} within 90 days (closest value prior to surgery start time), mean or median
    - Hemoglobin within 30 days (closest value prior to surgery start time), mean or median
3. Operating room
    - Laparoscopic procedure, #/%
    - Rectal procedure, #/%
    - Length of surgery, minutes, mean or median
    - Total intravenous (IV) fentanyl, mcg, mean or median
    - Received epidural or spinal, #/%
    - Intraoperative crystalloid IV fluids, mL (sum of normal saline, lactated ringers, and D5W), mean or median
4. Postoperative opioid
    - Day of surgery (DOS): oral morphine equivalent, mg, mean or median
    - Postoperative day (POD) 1: oral morphine equivalent, mg, mean or median
    - Postoperative day (POD) 2: oral morphine equivalent, mg, mean or median
5. Postoperative multimodal analgesia
    - Day of surgery
        - Oral acetaminophen, #/% of patients receiving any amount
        - IV acetaminophen, #/% of patients receiving any amount
        - IV or oral nonsteroidal anti-inflammatory drug (NSAID), #/% of patients receiving any amount
    - Postoperative day 1:
        - Oral acetaminophen, #/% of patients receiving any amount
        - IV acetaminophen, #/% of patients receiving any amount
        - IV or oral NSAID, #/% of patients receiving any amount
        - Multimodal agents, # of agents received, mean or median
    - Postoperative day 2:
        - Oral acetaminophen, #/% of patients receiving any amount
        - IV acetaminophen, #/% of patients receiving any amount
        - IV or oral NSAID, #/% of patients receiving any amount

- Multimodal agents, # of types of agents received, mean/median

- Postoperative day 3:
  - Oral acetaminophen, #/% of patients receiving any amount
  - IV acetaminophen, #/% of patients receiving any amount
- Total IV acetaminophen (DOS through POD 3), mg, mean or median
- Total oral acetaminophen (DOS through POD 3), mg, mean or median

6. Postoperative fluid
   - Day of surgery
     - Net input and output, mL, mean or median
     - Total input, mL, mean or median
   - · Postoperative day 1:
     - Net input and output, mL, mean or median
     - Total input, mL, mean or median
7. Postoperative patient experience
   - Postoperative patient-controlled analgesia use, hours, mean or median
   - Peak pain score on day of surgery, mean or median
   - Peak pain score on postoperative day 1, mean or median
   - Peak pain score on postoperative day 2, mean or median
   - Postoperative urinary straight catheterization on day of surgery, #/%
   - Postoperative urinary straight catheterization on postoperative day 1, #/%·
   - Postoperative insertion of urinary catheter (insertion occurs more than 2 hours after surgery stop or removal in operating room), #/%
   - Postoperative duration of urinary catheter, hours, mean or median
   - Time to first flatus or bowel movement, hours, mean or median
   - Postoperative duration of nasogastric tube, hours, mean or median
   - Postoperative insertion of nasogastric tube (insertion occurs more than 2 hours after surgery stop or removal in operating room), #/%
   - Delta creatinine (peak value within 72 hours postoperatively minus preoperative value), mean or median
   - Postoperative length of stay, days, mean or median
   - Discharge opioids (oral morphine equivalents) prescribed, mg, mean or median

Unless specified otherwise, units of measurement for continuous variables are mg and categorical end points are yes/no.

Tableau Visualization

Overview

The Datamart dashboard was designed to incorporate a user-friendly interface, allowing users to easily generate customizable reports. It offers tools to filter data by patient population, time frames, and specific metrics, enabling a focused analysis of ERAS outcomes (e.g., compare before and after implementation outcomes, track changes in protocols over specific time frames).

The user is first presented with a screening interface where patient population, time frames, and age parameters can be defined (Figure 1). By hovering over the colored square, the user can also review certain patient and surgery-specific information, such as associated *ICD-10* PCS, surgeon, and date of surgery (Figure 2). Patients can be deselected from the population if, for example, erroneous classification of an emergency trauma patient as an elective case.

Figure 1. User interface (screening). User interface depicting screening options. The user can select time frames, patient age range, and surgical population. The user is presented with the patient name, arrival and discharge date and time, patient age on day of surgery, qualifier for inclusion in the Datamart (*International Classification of Diseases, Tenth Revision* [*ICD-10*] *Procedure Coding System* [PCS] and/or surgical case procedure), and qualifier description. Patients can be selected or deselected on the right. Hovering over the colored square yields additional screening data (see Figure 2).

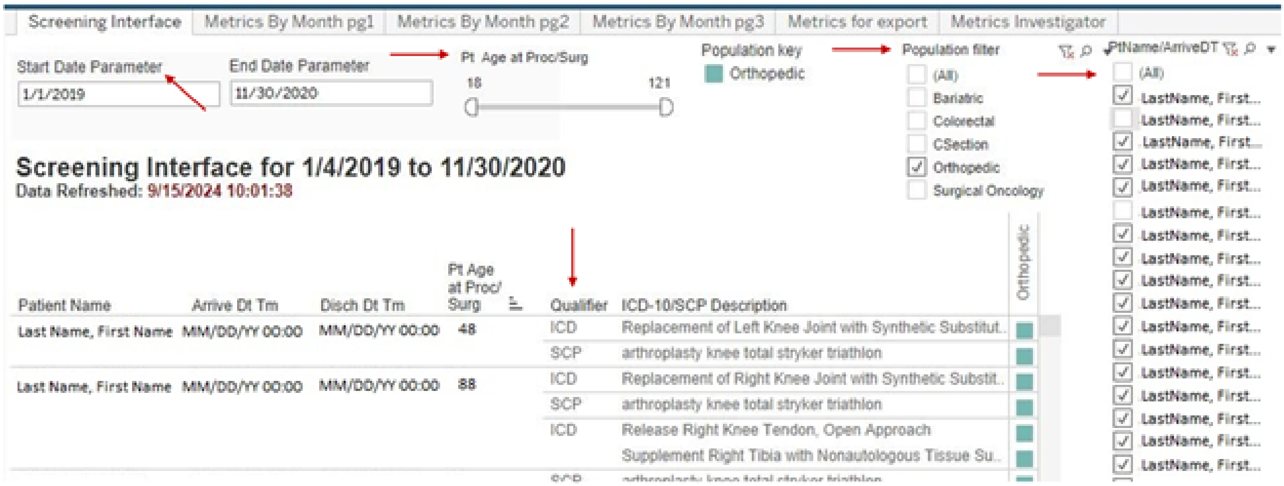
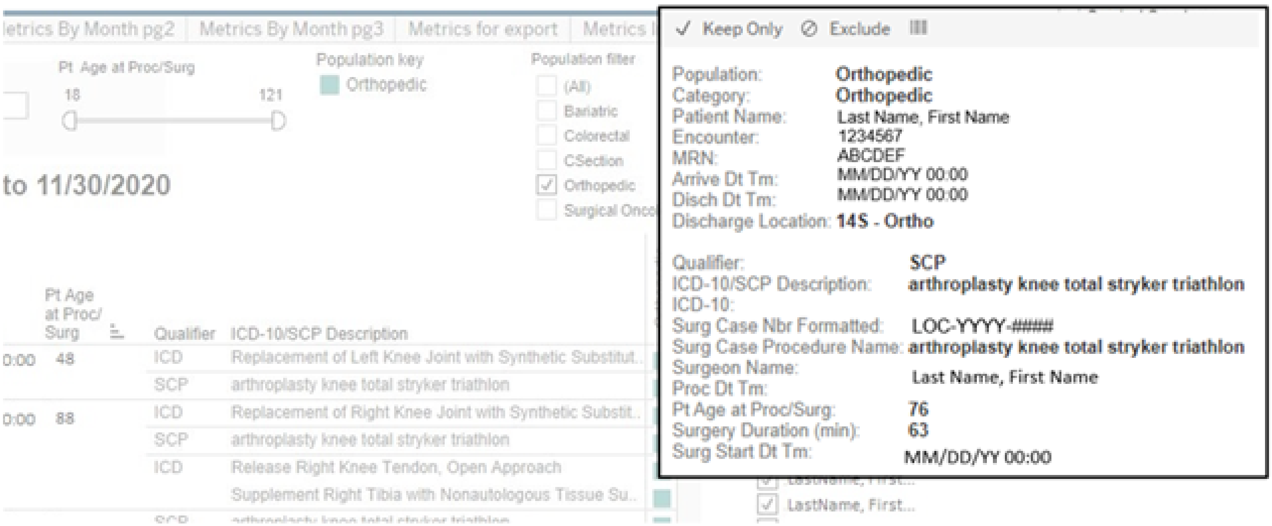
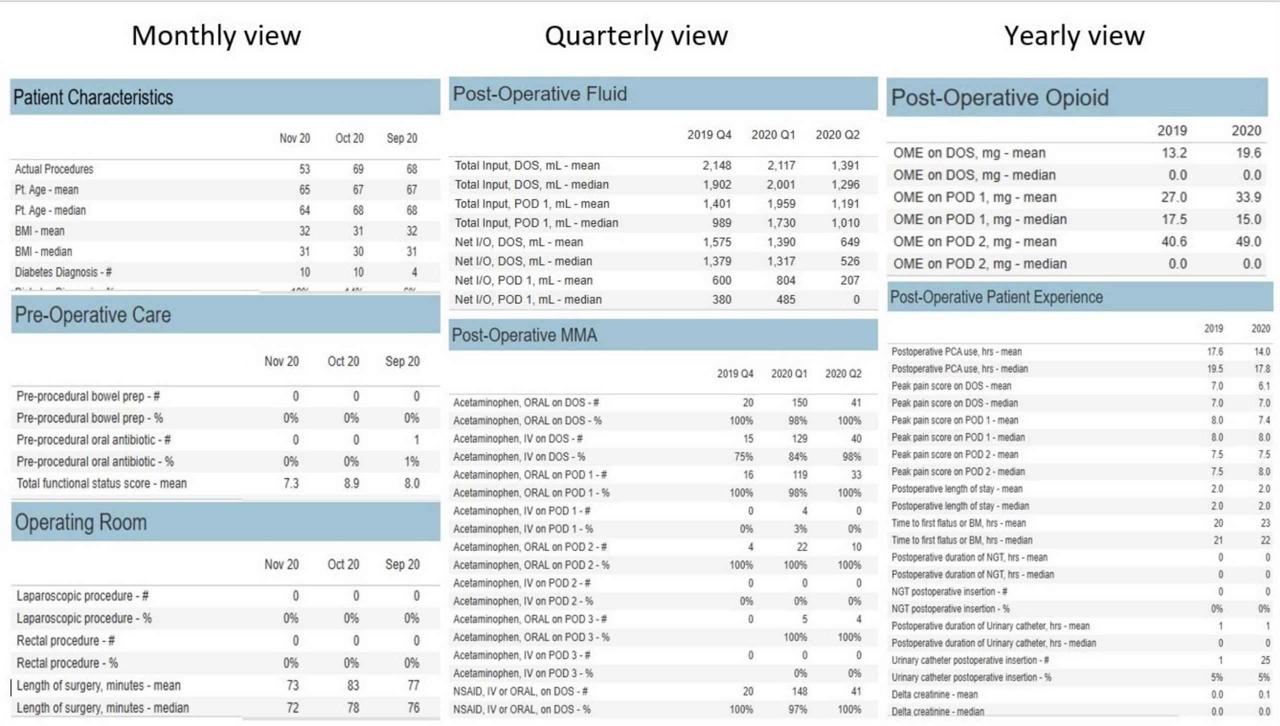


Figure 2. User interface (related screening). Additional screening information is available to the user. Hovering over the colored square next to the *International Classification of Diseases, Tenth Revision* (*ICD 10*)/surgical case procedure (SCP) description “pops out” information related to the selected case. This includes discharge location, surgical case number, surgeon name, and length of surgery.



On subsequent screens, the user can determine the metrics for analysis and how to view the data (Figure 3). Filters can be customized to focus on specific metrics or patient outcomes. For example, to focus on total acetaminophen use, one can view more details, such as “total IV acetaminophen on post-op day 1 (mean or median values).” Individual elements, such as opioid administration or fluid management details, can also be isolated for deeper analysis. Additionally, data can be displayed in weekly, monthly, quarterly, or yearly summaries and saved and exported as tables.

Figure 3. User interface (metrics). After procedures have been selected for review, the user can determine the metrics for analysis. Metrics can be viewed weekly, monthly, quarterly, or annually. Individual elements, such as opioid administration or fluid management details, can also be isolated for deeper analysis. DOS: day of surgery; IV: intravenous;nasogastric tube; MMA: multimodal analgesia; NGT: NSAID: nonsteroidal anti-inflammatory drug; OME: oral morphine equivalent; PCA: patient-controlled analgesia; POD: postoperative day.



Monthly view

Patient Characteristics

	Nov 20	Oct 20	Sep 20
Actual Procedures	53	69	68
Pt. Age - mean	65	67	67
Pt. Age - median	64	68	68
BMI - mean	32	31	32
BMI - median	31	30	31
Diabetes Diagnosis - #	10	10	4

Pre-Operative Care

	Nov 20	Oct 20	Sep 20
Pre-procedural bowel prep - #	0	0	0
Pre-procedural bowel prep - %	0%	0%	0%
Pre-procedural oral antibiotic - #	0	0	1
Pre-procedural oral antibiotic - %	0%	0%	1%
Total functional status score - mean	7.3	8.9	8.0

Operating Room

	Nov 20	Oct 20	Sep 20
Laparoscopic procedure - #	0	0	0
Laparoscopic procedure - %	0%	0%	0%
Rectal procedure - #	0	0	0
Rectal procedure - %	0%	0%	0%
Length of surgery, minutes - mean	73	83	77
Length of surgery, minutes - median	72	78	76

Quarterly view

Post-Operative Fluid

	2019 Q4	2020 Q1	2020 Q2
Total Input, DOS, mL - mean	2,148	2,117	1,391
Total Input, DOS, mL - median	1,902	2,001	1,296
Total Input, POD 1, mL - mean	1,401	1,959	1,191
Total Input, POD 1, mL - median	989	1,730	1,010
Net I/O, DOS, mL - mean	1,575	1,390	649
Net I/O, DOS, mL - median	1,379	1,317	526
Net I/O, POD 1, mL - mean	600	804	207
Net I/O, POD 1, mL - median	380	485	0

Post-Operative MMA

	2019 Q4	2020 Q1	2020 Q2
Acetaminophen, ORAL on DOS - #	20	150	41
Acetaminophen, ORAL on DOS - %	100%	98%	100%
Acetaminophen, IV on DOS - #	15	129	40
Acetaminophen, IV on DOS - %	75%	84%	98%
Acetaminophen, ORAL on POD 1 - #	16	119	33
Acetaminophen, ORAL on POD 1 - %	100%	98%	100%
Acetaminophen, IV on POD 1 - #	0	4	0
Acetaminophen, IV on POD 1 - %	0%	3%	0%
Acetaminophen, ORAL on POD 2 - #	4	22	10
Acetaminophen, ORAL on POD 2 - %	100%	100%	100%
Acetaminophen, IV on POD 2 - #	0	0	0
Acetaminophen, IV on POD 2 - %	0%	0%	0%
Acetaminophen, ORAL on POD 3 - #	0	5	4
Acetaminophen, ORAL on POD 3 - %		100%	100%
Acetaminophen, IV on POD 3 - #	0	0	0
Acetaminophen, IV on POD 3 - %		0%	0%
NSAID, IV or ORAL, on DOS - #	20	148	41
NSAID, IV or ORAL, on DOS - %	100%	97%	100%

Yearly view

Post-Operative Opioid

	2019	2020
OME on DOS, mg - mean	13.2	19.6
OME on DOS, mg - median	0.0	0.0
OME on POD 1, mg - mean	27.0	33.9
OME on POD 1, mg - median	17.5	15.0
OME on POD 2, mg - mean	40.6	49.0
OME on POD 2, mg - median	0.0	0.0

Post-Operative Patient Experience

	2019	2020
Postoperative PCA use, hrs - mean	17.6	14.0
Postoperative PCA use, hrs - median	19.5	17.8
Peak pain score on DOS - mean	7.0	6.1
Peak pain score on DOS - median	7.0	7.0
Peak pain score on POD 1 - mean	8.0	7.4
Peak pain score on POD 1 - median	8.0	8.0
Peak pain score on POD 2 - mean	7.5	7.5
Peak pain score on POD 2 - median	7.5	8.0
Postoperative length of stay - mean	2.0	2.0
Postoperative length of stay - median	2.0	2.0
Time to first flatus or BM, hrs - mean	20	23
Time to first flatus or BM, hrs - median	21	22
Postoperative duration of NGT, hrs - mean	0	0
Postoperative duration of NGT, hrs - median	0	0
NGT postoperative insertion - #	0	0
NGT postoperative insertion - %	0%	0%
Postoperative duration of Urinary catheter, hrs - mean	1	1
Postoperative duration of Urinary catheter, hrs - median	0	0
Urinary catheter postoperative insertion - #	1	25
Urinary catheter postoperative insertion - %	5%	5%
Delta creatinine - mean	0.0	0.1
Delta creatinine - median	0.0	0.0

For further in-depth analyses, for example, inferential statistical analysis, users can export patient-level raw data in either comma-separated value or Excel format. These data can be exported in a deidentified manner, safeguarding against the unintentional disclosure of protected health information when exiting the secure Datamart dashboard. By using these features, the Datamart permits users to analyze ERAS metrics with precision, adapt to different populations, and investigate trends and outcomes.

Use Case Examples: ERAS Total Joint Replacement and Surgical Oncology

Using total joint replacement as a use case example, we sought to compare patient care prior to the initiation of ERAS (pre-ERAS n=693 cases) and 12 months post-ERAS implementation (n=563 cases). Figures 4 and 5 show the impact of an ERAS program on pain using the data harvested from the Datamart, notably an apparent improvement in several pain-related metrics. Of note, the Tableau dashboard does not calculate or show error bars for continuous variables; the calculation of these requires export of the raw data for statistical analysis. Therefore, we opted to present mean and median, since if they are similar, it suggests that the data are normally distributed and the mean value is not inflated due to a few large outliers. Therefore, for most day-to-day QA purposes, the presentation of mean and median is sufficient, but for more rigorous quantitative analysis, for example, inferential statistics for hypothesis testing, the export of patient-level raw data allows that capability.

Figure 4. Enhanced recovery after surgery (ERAS) total joint replacement program: impact on pain. The Tableau dashboard does not display error bars for continuous variables; these require the export of patient-level raw data for statistical analysis, which is a capability of the Datamart. As ERAS pathways are primarily a quality assurance (QA) initiative, we report mean and median values, as their similarity suggests a roughly normal distribution without major outliers. DOS: day of surgery; IV: intravenous; NSAID: nonsteroidal anti-inflammatory drug; OME: oral morphine equivalent; POD: postoperative day.

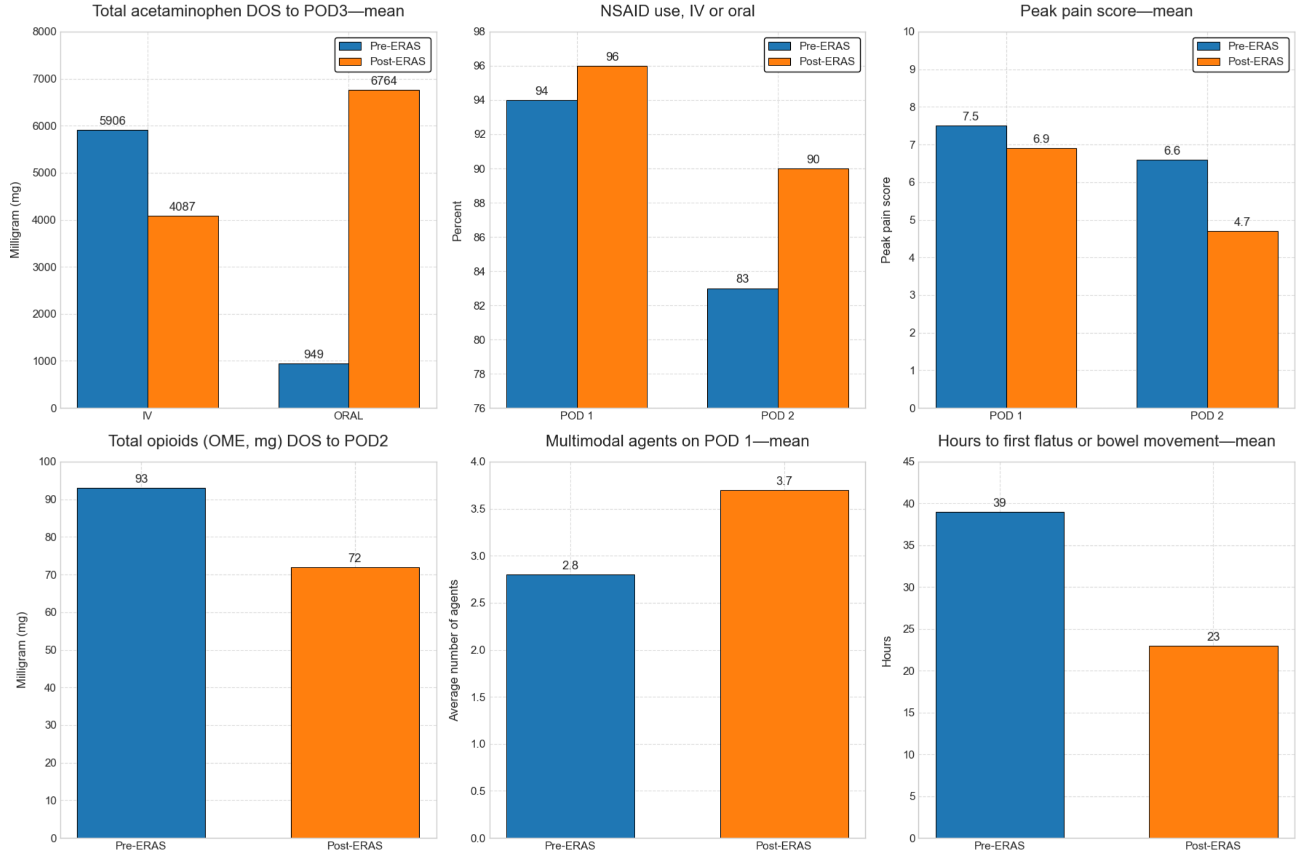
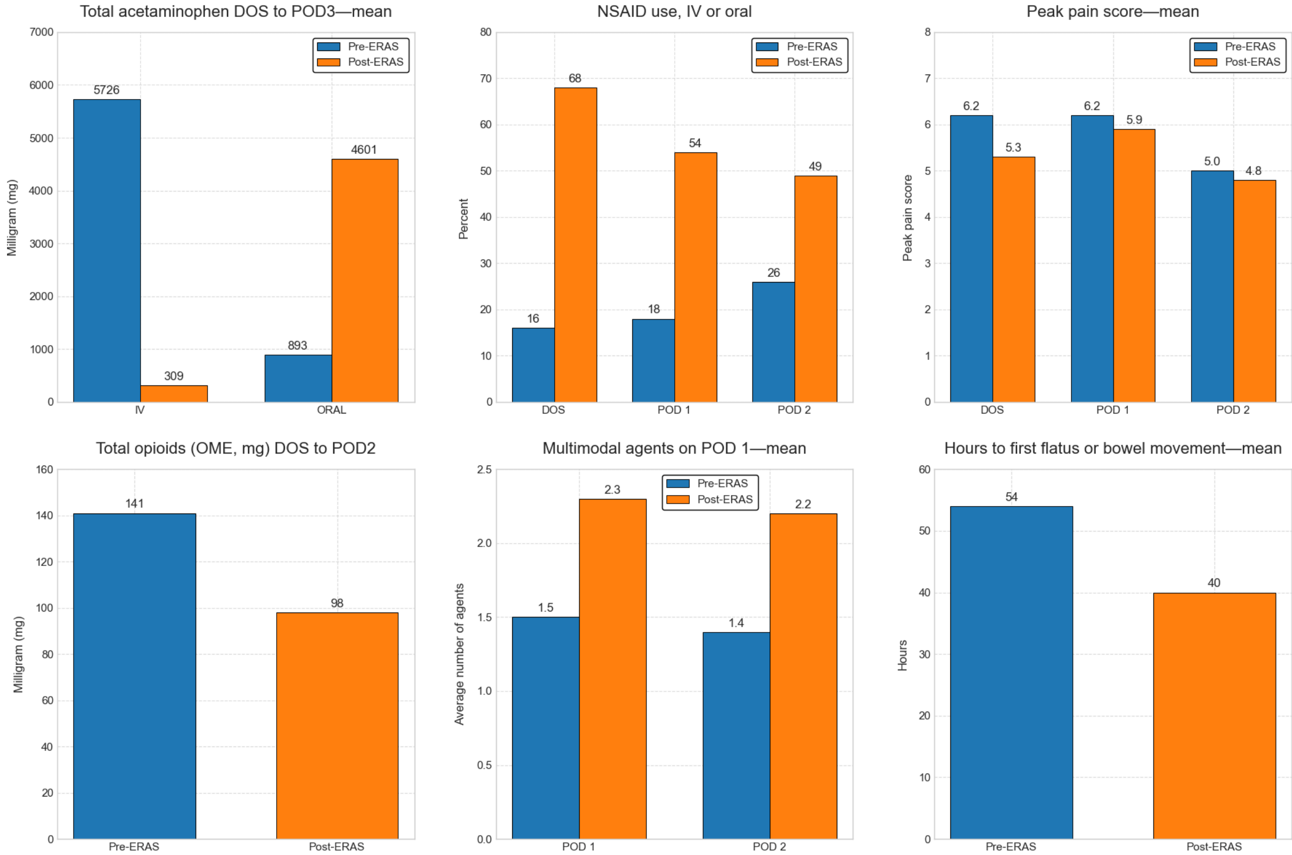


Figure 5. Enhanced recovery after surgery (ERAS) surgical oncology program: impact on pain. The Tableau dashboard does not display error bars for continuous variables; these require export of patient-level raw data for statistical analysis, which is a capability of the Datamart. As ERAS pathways are primarily a quality assurance (QA) initiative, we report mean and median values, as their similarity suggests a roughly normal distribution without major outliers. DOS: day of surgery; IV: intravenous; NSAID: nonsteroidal anti-inflammatory drug; OME: oral morphine equivalent; POD: postoperative day.



We noticed similar findings for our Surgical Oncology ERAS program (Figure 5).

Discussion

Principal Results

The novel Datamart dashboard has been a transformative tool at our hospital by enabling continuous analysis of ERAS programs and patient outcomes. By automating data extraction from the EMR, the Datamart eliminates the need for manual data abstraction, likely reducing errors and improving data accuracy. This centralized system currently serves as the primary data resource for our institutional ERAS programs, streamlining the monitoring of bundle elements and outcomes at both the individual and aggregate levels. Successful adoption required early stakeholder engagement, quarterly dashboard reviews, and review at quarterly stakeholder meetings. Barriers included initial unfamiliarity with Tableau capabilities and competing clinical priorities. These were mitigated through targeted training and dashboard use in routine quality meetings.

Key features of the Datamart include categorized metrics across perioperative phases, tailored reporting capabilities, and the ability to export data. The level of granularity of the Datamart provides actionable insights into ERAS pathway efficacy, supporting evidence-based decision-making.

Limitations

Despite its many advantages, the Datamart has some limitations that warrant mention. As with any EMR-based form of data capture, it relies on accurate charting by clinicians, which is not infallible. For example, postoperative ambulation was not included in this iteration, as documentation practices are inconsistent. Moreover, patient-reported outcomes, such as satisfaction and functional recovery, are not currently documented in the EMR, which prevents their inclusion in the Datamart.

Accurately identifying patients on an ERAS pathway is another challenge. The Datamart relies on *ICD-10* PCS for surgical procedure identification, which can lead to the inadvertent inclusion of non-ERAS cases. For example, procedures such as hemorrhoidectomy (non-ERAS) and hemicolectomy (ERAS) share the same *ICD-10* PCS, requiring manual filtering to exclude nonrelevant cases. While procedures, such as total knee replacement or cesarean section, are easier to identify due to constrained coding options, consistent oversight is critical for accurate data classification.

An intentional 2-month lag in data import further limits real-time analysis. However, this delay ensures the accuracy of *ICD-10* coding and surgical procedure inclusion, contributing to the integrity of the dataset.

Comparison With Prior Work

Health care auditing has significantly evolved with the advent of digital dashboards, data warehouses, and interactive audit systems. Before widespread dashboard integration, many institutions relied on manual data abstraction and semiautomated systems. Manual abstraction is cumbersome and relies on human resources to extract data from the EMR to input into a repository for analysis (e.g., REDCap, Microsoft Excel) [7-9]. Semiautomated audits (e.g. IT-generated reports), which at our institution can take 9 months or longer, add to the resource burden. Some institutions combine EMR data extraction with administrative databases (e.g., NSQIP) to build a centralized reporting structure [2,10]. However, reliance on these reports poses challenges such as delayed access to real-time data, lack of customization, and risk of system downtimes.

For some institutions with ERAS pathways, third-party data warehouses (e.g., EIAS) may be an option [5]. While they offer additional storage and analysis options, third-party warehouses present issues of data security, higher costs, and limited control over information (see Table 1). In contrast, our in-house customizable Datamart can support diverse ERAS populations and continuous improvements through iterative refinements based on user feedback and technological advancements.

Table . Comparison between the Datamart and third-party options.

Feature	Datamart+Tableau	EIAS[a]	ACS[b] NSQIP[c]
Data control	Full institutional control over data	Vendor-controlled; limited flexibility	External benchmarking database
Customization	Highly customizable by procedure type, metrics, and time frame	Some customization possible, but limited	Standardized measures; less customizable
Real-time access	Near real-time (2-month lag for data integrity)	Typically delayed, relies on upload	Reports released quarterly or semi-annually
Cost	Internal development; no licensing fees	Requires subscription or license fees	Expensive participation and data access
Data security	Remains within institutional firewall; HIPAA[d]-compliant	Data housed externally; possible security concerns	Data deidentified but externally stored
Metric flexibility	Fully institution-defined (≥56 metrics currently used)	Limited to ERAS[e]-recommended fields	Fixed set of standard metrics
Scalability	Easily scalable across services and use cases	Limited by system design and vendor	Only covers specific surgeries (e.g., colectomy)

[a]EIAS: ERAS Interactive Audit System.

[b]ACS: American College of Surgeons.

[c]NSQIP: National Surgical Quality Improvement Program.

[d]HIPAA: Health Insurance Portability and Accountability Act.

[e]ERAS: enhanced recovery after surgery.

While we were unable to find reports in the literature of the use of a datamart with dashboard visualization tools to support ERAS programs, these tools have been described in other types of quality, clinical, and research programs. Institutionally developed data infrastructures have been used to aggregate disparate data sources into unified repositories to support research and clinical audits [11-13]. Interactive dashboards (e.g., Tableau and Qlik) have been implemented to visualize clinical performance and facilitate quality improvement efforts [14-16]. Thus, enabling data-driven feedback to stakeholders. The Datamart combines the filtering and query capabilities of a unified data repository with the visualization tools of interactive dashboards to improve patient outcomes, enhance adherence to protocols, improve interdisciplinary communication, and support decision-making.

Conclusions

The Datamart represents a significant advancement in ERAS program management. Unlike third-party systems, such as the EIAS or ACS NSQIP, the Datamart provides full institutional control over data, customizable metrics aligned with local priorities, and flexible reporting capabilities. Although third-party systems can be limited by fixed datasets, external data hosting, and reporting delays, the Datamart enables near real-time internal data access, tailored QA tracking, and the ability to refine metrics, as clinical needs evolve. This offers a robust, centralized, and automated approach to data monitoring and analysis without recurring licensing costs or dependence on vendor timelines. The ability to reduce manual abstraction, improve data accuracy, and evaluate intervention makes the Datamart a vital tool for enhancing perioperative care.

There are several promising directions for the enhancement and expansion of this system. Integrating predictive analytics, automated alerts, natural language processing, and machine learning into the Datamart could enable users to better anticipate complications and tailor interventions [11,12]. For example, predictive models could identify patients at high risk for ERAS noncompliance or adverse outcomes, allowing proactive adjustments to care plans. Automated alerts embedded within the dashboard could notify clinicians in real time when critical metrics fall outside expected ranges, supporting timely interventions and reducing preventable complications.

Expanding the dashboard to encompass additional surgical specialties and incorporating patient-reported outcomes, through structured EMR fields or digital surveys, into the Datamart could further enhance applicability and patient-centered care. These innovations would transform the Datamart from a retrospective monitoring tool into a dynamic, decision-support platform that drives continuous improvement in perioperative care.

As ERAS programs grow in scope and complexity, the need for scalable, adaptable solutions to implementing and monitoring evidence-based care and patient outcomes is increasingly evident. This novel Datamart dashboard addresses many of these challenges while providing a foundation for ongoing innovation. Although developed within Cerner, the Datamart framework is adaptable to other EMRs, provided discrete data fields are available. Implementation requires collaboration with nursing, anesthesiology and surgical services, institutional IT, along with Tableau or similar visualization tools. Institutions seeking to improve ERAS monitoring may consider adapting the framework described here, tailoring metric section and dashboard design to their local EMR environment and clinical priorities.

Data Availability

The datasets used are not available.

Authors' Contributions

Conceptualization: EB-G (lead), SMS (equal)
Data curation: EB
Formal analysis: AS (supporting), EB (lead), JN (supporting), SMS (supporting)
Funding acquisition: EB-G
Investigation: EB (equal), SMS (equal)
Methodology: EB (equal), EB-G (lead), SMS (equal)
Project administration: EB (supporting), EB-G (lead), SMS (equal)
Resources: EB
Software: EB
Supervision: EB-G
Validation: EB (equal), SMS (equal)
Visualization: EB-G (lead), SMS (supporting)
Writing – original draft: EB-G (equal), SO (equal), SMS (equal)
Writing – review & editing: AS (supporting), EB (supporting), EB-G (lead), JN (supporting), SMS (equal), SO (equal)

Conflicts of Interest

None declared.

References

1. Gramlich LM, Sheppard CE, Wasylak T, et al. Implementation of enhanced recovery after surgery: a strategy to transform surgical care across a health system. Implement Sci 2017 May 19;12(1):67. [doi: 10.1186/s13012-017-0597-5] [Medline: 28526041]
2. Thiele RH, Rea KM, Turrentine FE, et al. Standardization of care: impact of an enhanced recovery protocol on length of stay, complications, and direct costs after colorectal surgery. J Am Coll Surg 2015 Apr;220(4):430-443. [doi: 10.1016/j.jamcollsurg.2014.12.042] [Medline: 25797725]
3. Singh SM, Liverpool A, Romeiser JL, Thacker J, Gan TJ, Bennett-Guerrero E. Types of surgical patients enrolled in enhanced recovery after surgery (ERAS) programs in the USA. Perioper Med (Lond) 2021 Apr 27;10(1):12. [doi: 10.1186/s13741-021-00185-5] [Medline: 33902705]
4. Carey S, Hogan S. Failure in systems and culture: barriers that prevent implementation of evidence-based fasting times for patients in the acute care setting. JPEN J Parenter Enteral Nutr 2021 Jul;45(5):933-940. [doi: 10.1002/jpen.1961] [Medline: 32654214]
5. ERAS interactive audit system - EIAS by Encare. Encare Enhancing Eras. URL: https://encare.net/products-and-services/eras-interactive-audit-system-eias/ [accessed 2024-12-12]
6. Langnas EM, Matthay ZA, Lin A, et al. Enhanced recovery after surgery protocol and postoperative opioid prescribing for cesarean delivery: an interrupted time series analysis. Perioper Med (Lond) 2021 Nov 15;10(1):38. [doi: 10.1186/s13741-021-00209-0] [Medline: 34775985]
7. Carter J. Fast-track surgery in gynaecology and gynaecologic oncology: a review of a rolling clinical audit. ISRN Surg 2012;2012:368014. [doi: 10.5402/2012/368014] [Medline: 23320193]
8. Archibald LH, Ott MJ, Gale CM, Zhang J, Peters MS, Stroud GK. Enhanced recovery after colon surgery in a community hospital system. Dis Colon Rectum 2011 Jul;54(7):840-845. [doi: 10.1007/DCR.0b013e31821645bd] [Medline: 21654251]

9. Sinha R, Verma N, Bana R, Kalidindi N, Sampurna S, Mohanty GS. Intra- and post-operative outcomes in benign gynaecologic surgeries before and after the implementation of enhanced recovery after surgery protocols: a comparison. J Minim Access Surg 2023;19(1):112-119. [doi: 10.4103/jmas.jmas_42_22] [Medline: 36722536]
10. Ripollés-Melchor J, Abad-Motos A, Díez-Remesal Y, et al. Association between use of enhanced recovery after surgery protocol and postoperative complications in total hip and knee arthroplasty in the Postoperative Outcomes Within Enhanced Recovery After Surgery Protocol in Elective Total Hip and Knee Arthroplasty Study (POWER2). JAMA Surg 2020 Apr 1;155(4):e196024. [doi: 10.1001/jamasurg.2019.6024] [Medline: 32049352]
11. Marazzi F, Tagliaferri L, Masiello V, et al. GENERATOR Breast DataMart—the novel breast cancer data discovery system for research and monitoring: preliminary results and future perspectives. J Pers Med 2021 Jan 22;11(2):65. [doi: 10.3390/jpm11020065] [Medline: 33498985]
12. D'Amario D, Laborante R, Delvinioti A, et al. GENERATOR HEART FAILURE DataMart: an integrated framework for heart failure research. Front Cardiovasc Med 2023;10:1104699. [doi: 10.3389/fcvm.2023.1104699] [Medline: 37034335]
13. Hurst JH, Liu Y, Maxson PJ, Permar SR, Boulware LE, Goldstein BA. Development of an electronic health records datamart to support clinical and population health research. J Clin Transl Sci 2020 Jun 23;5(1):e13. [doi: 10.1017/cts.2020.499] [Medline: 33948239]
14. Lechner C, Rumpler M, Dorley MC, Li Y, Ingram A, Fryman H. Developing an online dashboard to visualize performance data—Tennessee newborn screening experience. Int J Neonatal Screen 2022 Sep 2;8(3):49. [doi: 10.3390/ijns8030049] [Medline: 36135349]
15. Dang QCL, Román E, Donner K, et al. Implementation and impact of a surgical dashboard on pediatric tonsillectomy outcomes: a quality improvement study. Laryngoscope Investig Otolaryngol 2024 Oct;9(5):e1315. [doi: 10.1002/lio2.1315] [Medline: 39281202]
16. Alhmoud B, Melley D, Khan N, Bonicci T, Patel R, Banerjee A. Evaluating a novel, integrative dashboard for health professionals' performance in managing deteriorating patients: a quality improvement project. BMJ Open Qual 2022 Dec;11(4):e002033. [doi: 10.1136/bmjoq-2022-002033] [Medline: 36588306]

Abbreviations

ACS: American College of Surgeons
EIAS: ERAS Interactive Audit System
EMR: electronic medical record
ERAS: Enhanced Recovery after Surgery
***ICD-10*:** *International Classification of Diseases, Tenth Revision*
NSQIP: National Surgical Quality Improvement Program
PCS: Procedure Coding System
POD: postoperative day
QA: quality assurance
REDCap: Research Electronic Data Capture

Edited by N Rohatgi; submitted 15.Aug.2025; peer-reviewed by AR Cochran, C McLane; revised version received 15.Dec.2025; accepted 18.Dec.2025; published 11.Feb.2026.

Please cite as:
Singh SM, Oster S, Bolze E, Sasson A, Nicholson J, Bennett-Guerrero E
A Novel Customizable Datamart and Tableau Dashboard to Monitor Multiple Enhanced Recovery After Surgery Programs: Development and Validation Study
JMIR Perioper Med 2026;9:e82472
URL: https://periop.jmir.org/2026/1/e82472
doi:10.2196/82472

Survival Prediction in Patients With Bladder Cancer Undergoing Radical Cystectomy Using a Machine Learning Algorithm: Retrospective Single-Center Study

Francesco Andrea Causio[1,2], MD; Vittorio De Vita[1,2], MD; Andrea Nappi[1,3], BSc; Melissa Sawaya[4], PhD; Bernardo Rocco[5,6], MD, PhD; Nazario Foschi[5,6], MD; Giuseppe Maioriello[5,6], MD; Pierluigi Russo[1,5,6,7], MD

[1]Italian Society for Artificial Intelligence in Medicine (SIIAM - Società Italiana Intelligenza Artificiale in Medicina), Rome, Italy
[2]University Department of Life Sciences and Public Health, Section of Hygiene, Università Cattolica del Sacro Cuore, Largo F Vito 1, Rome, Italy
[3]Computer Science Department, University of Twente, Twente, The Netherlands
[4]Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, Villejuif, France
[5]University Department of Urology, Università Cattolica del Sacro Cuore, Rome, Italy
[6]University Department Medicine and Translational Surgery, Università Cattolica del Sacro Cuore, Rome, Italy
[7]Department of Life Science, Health, and Health Professions, Università Degli Studi Link, Rome, Italy

Corresponding Author:
Vittorio De Vita, MD
Italian Society for Artificial Intelligence in Medicine (SIIAM - Società Italiana Intelligenza Artificiale in Medicina), Rome, Italy

Abstract

Background: Traditional statistical models often fail to capture the complex dynamics influencing survival outcomes in patients with bladder cancer after radical cystectomy, a procedure where approximately 50% of patients develop metastases within 2 years. The integration of artificial intelligence (AI) offers a promising avenue for enhancing prognostic accuracy and personalizing treatment strategies.

Objective: This study aimed to develop and evaluate a machine learning algorithm for predicting disease-free survival (DFS), overall survival (OS), and the cause of death in patients with bladder cancer undergoing cystectomy, using a comprehensive dataset of clinical and pathological variables.

Methods: Retrospective data of 370 patients with bladder cancer who underwent radical cystectomy at Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy, were collected. The dataset comprised 20 input variables, encompassing demographics, tumor characteristics, treatment variables, and inflammatory markers. For specific analyses and models, we used patient subcohorts. The CatBoost algorithm was used for regression tasks (DFS in 346 patients, OS in 347 patients) and a binary classification task (tumor-related death in 312 patients). Model performance was assessed using mean absolute error (MAE) for regression and F_1-score for classification, prioritizing a minimum recall of 75% for tumor-related deaths. Five-fold cross-validation and Shapley additive explanations (SHAP) values were used to ensure robustness and interpretability.

Results: For DFS prediction, the CatBoost model achieved an MAE of 18.68 months, with clinical tumor stage and pathological tumor classification identified as the most influential predictors. OS prediction yielded an MAE of 17.2 months, which improved to 14.6 months after feature filtering, where tumor classification and the systemic immune-inflammation index (SII) were most impactful. For tumor-related death classification, the model achieved a recall of 78.6% and an F_1-score of 0.44 for the positive class (tumor-related deaths), correctly identifying 11 of 14 cases. Bladder tumor position was the most influential feature for cause-of-death prediction.

Conclusions: The developed machine learning algorithm demonstrates promising accuracy in predicting survival and the cause of death in patients with bladder cancer after cystectomy. The key predictors include clinical and pathological tumor staging, systemic inflammation (SII), and bladder tumor position. These findings highlight the potential of AI in providing clinicians with an objective, data-driven tool to improve personalized prognostic assessment and guide clinical decision-making.

(JMIR Perioper Med 2026;9:e86666) doi:10.2196/86666

KEYWORDS

cystectomy; disease-free survival; artificial intelligence; neoplasm staging; retrospective studies; urinary bladder neoplasms; clinical decision-making; machine learning; statistical models

Introduction

In the evolving landscape of health care, the integration of artificial intelligence (AI) into clinical decision-making has gained significant momentum, particularly in the realm of oncology [1,2]. With advancements in machine learning techniques, health care professionals are increasingly harnessing the power of AI to enhance diagnosis, prognosis, and treatment planning. The exponential growth of digital health care data, including electronic health records, medical imaging, genomic data, and real-time patient monitoring, has fueled the development of predictive algorithms [1,3].

The field of urology is complex: cancerous conditions benefit from the leverage of additional data sources and decision-making algorithms that allow physicians to plan treatment while considering several complex factors. Urological cancers, including prostate, bladder, and renal cancers, place a considerable burden on health care systems worldwide [4]. These malignancies often require complex management involving early diagnosis, accurate staging, and personalized treatment strategies to optimize patient outcomes. Traditional methods of assessing prognosis rely heavily on statistical models that may not capture the multifaceted nature of cancer behavior and patient responses to treatment. Conventional regression statistics often fail to provide the depth of analysis required to address the complexities of cancer management. In contrast, AI techniques, such as artificial neural networks, Bayesian networks, and neuro-fuzzy modeling systems, offer innovative approaches to constructing data-driven models that can adapt to the heterogeneous nature of cancer [5].

The potential of AI in predicting patient outcomes is particularly evident in its ability to analyze large datasets without the constraints of predetermined statistical distributions. By leveraging retrospective data, we can develop algorithms that not only identify patterns and correlations but also provide insights into individual patient behavior. This capability is crucial for clinicians who face the challenge of tailoring treatment plans to the unique characteristics of each patient. In the context of mortality and postoperative survival, the application of AI can provide critical insights that enhance our understanding of patient outcomes following surgical interventions. The ability to predict which patients are at higher risk of complications or recurrence can lead to more informed clinical decisions, ultimately improving the quality of care [6]. For instance, machine learning algorithms can analyze a multitude of variables, including clinical, pathological, and demographic factors, to generate individualized risk profiles that guide treatment strategies and follow-up plans [7]. Recent urological research has shown that combining hematological inflammation indexes with machine learning algorithms can improve the prediction of surgical outcomes, as demonstrated in patients who underwent urethroplasty [8].

In this study, we focus specifically on the training of an AI algorithm using retrospective data collected from patients diagnosed with bladder cancer who underwent radical cystectomy. Patients with localized muscle-invasive or recurrent non–muscle-invasive bladder cancer benefit most from radical cystectomy, which may be preceded by neoadjuvant chemotherapy in selected cases, in terms of local disease control. Even with sufficient local control achieved through cystectomy, approximately 50% of patients develop metastases within 2 years and may ultimately die from the disease. This is likely due to the existence of regional or distant microscopic metastatic disease at the time of surgery [9]. The proposed methodology will involve the comprehensive examination of variables associated with patient demographics, tumor characteristics, treatment modalities, and postoperative outcomes. Using machine learning techniques, we aim to identify key predictors of mortality and postoperative survival, ultimately constructing a predictive model with potential relevance for clinical decision-making.

Methods

Study Design

We collected retrospective data on patients with high-risk and very high-risk non–muscle-invasive bladder cancer and muscle-invasive bladder cancer who underwent radical cystectomy at Fondazione Policlinico Universitario Agostino Gemelli IRCCS in Rome, Italy. The dataset included data on various clinical and pathological variables from 370 patients.

Ethical Considerations

Ethical approval was obtained from the institutional ethical review board (protocol number 676 - 02). As primary consent for data collection covered secondary analyses, additional consent was not required for this study. The data used in this study were anonymized, and no compensation was provided to patients.

Data Collection and Preprocessing

Clinical and pathological data were extracted from medical records, including demographic, lifestyle, tumor, treatment, and laboratory variables. The dataset was split into three outcome-specific subsets to maximize the usable sample for each task:

1. DFS dataset: predicting disease-free survival (DFS) in months (346 patients in total).
2. OS dataset: predicting overall survival (OS) in months (347 patients in total).
3. Death cause dataset: for classification purposes, the classes “death from other causes” and “alive” were merged into a single negative class to create a binary variable. Therefore, the cause is defined as either “no” or “yes,” depending on whether it was tumor related (312 patients in total).

The variables included in the dataset are detailed in Table 1. All categorical variables were cast as strings to allow native handling by CatBoost (version 1.2.8; Yandex).

A total of 20 input variables were selected for model development.

Table . Variables included in the study

Variable (English)	Description	Data type	Input/output
Patient demographics and lifestyle			
AGE	Age (years)	Numerical	Input
BMI	Body mass index (kg/m^2)	Numerical	Input
SEX	Biological sex (0: man, 1: woman)	Categorical	Input
SMOKE	Smoker (0: no, 1: yes)	Categorical	Input
Patient medical history			
DM	Patient has diabetes mellitus (0: no, 1: yes)	Categorical	Input
PRIOR SURGERY	Patient had previously undergone surgery in the abdominal area (0: no, 1: yes)	Categorical	Input
PRIOR RADIOTHERAPY	Patient had previously received radiotherapy in the abdominal area (0: no, 1: yes)	Categorical	Input
PRIOR SYSTEMIC CHEMOTHERAPY	Patient had previously received systemic chemotherapy (0: no, 1: yes)	Categorical	Input
Tumor characteristic			
BLADDER TUMOR POSITION	Identifier of tumor position (0: intertrigonal zone, 1: right periosteal, 2: left periosteal, 3: dome, 4: posterior wall, 5: right lateral wall, 6: left lateral wall, 7: prostatic urethra, 8: anterior wall, 9: entire bladder, 10: bladder base)	Categorical	Input
TUMOR DIMENSION	Tumor dimension (cm)	Numerical	Input
PRE-HYDRONEPHROSIS	Hydronephrosis (0: no, 1: right hydronephrosis, 2: left hydronephrosis, 3: bilateral hydronephrosis)	Categorical	Input
H.E. TURV	Histological examination for transurethral resection of the bladder (0: localized to mucosa +/– submucosa multirecurrent, 1: muscle-invasive, 2: squamous)	Categorical	Input
LVI	Lymphovascular invasion (0: absent, 1: present)	Categorical	Input
CTS	Clinical tumor stage (0: cTa, 1: cTis, 2: cT1, 3: cT2, 4: cT3, 5: cT4)	Categorical	Input
TC	Tumor classification (1: T0, 2: Ta, 3: Tis, 4: T1, 5: T2a, 6: T2b, 7: T3a, 8: T3b, 9: T4a, 10: T4b)	Categorical	Input
Inflammatory and immune marker			
SII	Systemic immune-inflammation index (decimals)	Numerical	Input
Treatment and outcome			
UD	Urinary diversion type (0: Bricker ileal conduit, 1: ureterocutaneostomy, 2: vesicoileal pouch)	Categorical	Input
RECURRENCE	Tumor recurrence (0: no, 1: yes)	Categorical	Input
DFS	Disease-free survival after treatment (in months)	Numerical	Output

Variable (English)	Description	Data type	Input/output
OS	Overall survival: time from diagnosis/treatment start to death from any cause (in months)	Numerical	Output
DEATH CAUSE	Cause of death (X: alive, 1: other, 2: cancer); later merged (0: alive + other, 1: cancer)	Categorical	Output

Machine Learning Models

To predict clinical outcomes, we used the CatBoost algorithm for both regression (DFS and OS) and classification (cause of death) tasks, as it is effective for small and structured datasets.

For DFS and OS, we applied CatBoostRegressor models. For predicting the cause of death, we used the CatBoostClassifier, with the binary outcome of death being tumor related or not.

DFS and OS Model Evaluation

For the regression tasks (DFS and OS), we evaluated model performance using mean absolute error (MAE) to quantify the average prediction error in months.

Cause-of-Death Classification Model

For the classification task (cause of death), we evaluated performance using the F_1-score. The F_1-score is a single metric that balances precision and recall, particularly useful in cases of imbalanced classes where the positive class is of primary interest. For class 1 (tumor-related deaths), it was calculated as the harmonic mean of precision and recall:

(1)F1=2∗(Precision∗Recall)/(Precision+Recall)

Precision is the proportion of correctly identified positive predictions among all positive predictions, and recall is the proportion of correctly identified positive predictions among all actual positives.

(2)Precision=TP/(TP+FP)

(3)Recall=TP/(TP+FN)

Confusion matrices were used to examine prediction distributions, and probability thresholds were adjusted to optimize recall while limiting false positives. To account for class imbalance in the classification task, we applied custom class weights. We adjusted the decision threshold, aiming for a minimum recall of 75% to ensure that most tumor-related deaths were accurately identified and classified.

Cross-Validation and Hyperparameter Tuning

All models were trained and evaluated using 5-fold cross-validation to ensure generalizability and reduce the risk of overfitting, especially given the relatively small dataset. In addition, we applied early stopping with a patience range of 30 to 50 rounds, allowing the model to terminate training once performance ceased to improve on the validation fold.

To enhance the interpretability and transparency of the developed machine learning models, we used violin plots and Shapley additive explanations (SHAP) scatterplots to investigate the impact of variables on the prediction of the results. Violin plots show the effect of each variable on the results, both in terms of direction (favorable or unfavorable) and intensity. The SHAP scatterplot assigns an importance value to each feature for a particular prediction. For each patient, the SHAP values revealed the specific features driving the predicted risk of tumor death. By aggregating the SHAP values across the entire cohort, the overall impact and importance of each clinical and pathological variable on the model's outcome predictions were determined. This enabled the identification of the most significant factors influencing the outcomes of patients with bladder cancer after cystectomy.

This paper presents only the most significant results of the analysis. The complete analysis is available online for open access [10].

Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis Guidelines for AI

To enhance the transparency, interpretability, and reproducibility of our machine learning-based prediction models, this study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement, specifically considering the extensions for AI (TRIPOD+AI). The TRIPOD+AI guidelines provide a standardized framework for reporting studies that develop or validate prediction models, ensuring that sufficient detail is provided for critical appraisal and replication by other researchers. By following these guidelines, we aim to clearly articulate the study design, data characteristics, model development process, and performance evaluation, thereby contributing to the responsible and rigorous application of AI in medical research (Table 2).

Table . Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis checklist for reporting studies involving artificial intelligence (TRIPOD+AI).

TRIPOD+AI item	Description of reporting in this study
1. Title	• Survival Prediction in Patients with Bladder Cancer Undergoing Radical Cystectomy Using a Machine Learning Algorithm: Retrospective Single-Center Study
2. Abstract	• The abstract summarizes the study's objectives, methods, key findings, and conclusions.
3. Introduction - background	• The introduction will establish the clinical context of bladder cancer, the prognostic challenges, and the rationale for using machine learning.
4. Methods - participants	
4a. Eligibility criteria	• Patients who underwent radical cystectomy for bladder cancer, with available data for selected variables
4b. Settings and locations	• Data collected retrospectively from a single institution: Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy
4c. Source of data	• Patient medical records
5. Data acquisition method	• Retrospective data extraction into a spreadsheet
6. Methods - outcome	
6a. Definition of outcomes	• Disease-free survival (DFS): time in months from treatment to recurrence or death (event) or last follow-up (censored) • Overall survival (OS): time in months from diagnosis/treatment start to death from any cause (event) or last follow-up (censored) • DEATH CAUSE: binary classification (0: did not die from the tumor, 1: died from the tumor)
6b. Outcome measurement	• DFS and OS were calculated from documented dates. • The cause of death was extracted from medical records and recategorized for binary classification.
7. Methods - predictors	

TRIPOD+AI item	Description of reporting in this study
7a. Definition of all predictors	• AGE: patient's age (years) • BMI (kg/m^2) • DM: patient has diabetes mellitus (0: no, 1: yes) • PRIOR SURGERY: patient had previously undergone surgery in the abdominal area (0: no, 1: yes) • PRIOR RADIOTHERAPY: patient had previously received radiotherapy in the abdominal area (0: no, 1: yes) • PRIOR SYSTEMIC CHEMOTHERAPY: patient had previously received systemic chemotherapy (0: no, 1: yes) • BLADDER TUMOR POSITION: identifier of tumor position (0: intertrigonal zone, 1: right periosteal, 2: left periosteal, 3: dome, 4: posterior wall, 5: right lateral wall, 6: left lateral wall, 7: prostatic urethra, 8: anterior wall, 9: entire bladder, 10: bladder base) • TUMOR DIMENSION: tumor dimension (cm) • PRE-HYDRONEPHROSIS: pretreatment hydronephrosis (0: no, 1: right hydronephrosis, 2: left hydronephrosis, 3: bilateral hydronephrosis) • SEX: biological sex (0: man, 1: woman) • SMOKE: patient smokes (0: no, 1: yes) • H.E. TURV: histological examination for transurethral resection of the bladder (0: localized to mucosa +/– submucosa multirecurrent, 1: muscle-invasive, 2: squamous) • SII: systemic immune-inflammation index (decimals) • UD: urinary diversion type (0: Bricker ileal conduit, 1: ureterocutaneostomy, 2: vesicoileal pouch) • LVI: lymphovascular invasion (0: absent, 1: present) • CTS: clinical tumor stage (0: cTa, 1: cTis, 2: cT1, 3: cT2, 4: cT3, 5: cT4) • TC: tumor classification (1: T0, 2: Ta, 3: Tis, 4: T1, 5: T2a, 6: T2b, 7: T3a, 8: T3b, 9: T4a, 10: T4b)
7b. Predictor measurement	• Predictors were measured clinically (eg, age and BMI), derived from patient history (eg, prior surgeries and smoking status), or derived from laboratory or pathology reports (eg, SII, tumor dimension, LVI, CTS, and TC).
8. Methods - sample size	
8a. Sample size determination	• The available retrospective data determined the sample size. No formal power calculation was performed due to the exploratory nature of the study and the limitations of the data.
9. Methods - data handling	
9a. Handling of missing data	• Rows with null values in specific critical variables ("TUMOR DIMENSION," "LVI," "TC," "H.E. TURV," "RECURRENCE," "DFS," "OS," "DEATH CAUSE") were removed. No imputation was performed.
9b. Data transformation	• Numerical variables were type-adjusted to int or float. Categorical variables were explicitly converted to category type. "DEATH CAUSE" was recategorized into a binary format.
10. Methods - model development	

TRIPOD+AI item	Description of reporting in this study
10a. Model type	• CatBoostRegressor (for DFS and OS) and CatBoostClassifier (for DEATH CAUSE).
10b. Candidate predictors	• All 17 selected independent variables were used as candidate predictors for each model, based on the relevant dataset (df1 [DFS], df2 [OST], df3 [DEATH CAUSE]).
10c. Handling of continuous predictors	• Continuous predictors (AGE, BMI, TUMOR DIMENSION, SII) were used directly by CatBoost, which handles them internally.
10d. Handling of categorical predictors	• Categorical predictors were identified and explicitly converted to string type before training. CatBoost natively handles categorical features without explicit one-hot encoding.
10e. Details of model fitting	• CatBoostRegressor (iterations=1000, learning_rate=0.05, depth=6, loss_function=”RMSE,” eval_metric=”MAE,” early_stopping_rounds=50, random_seed=42, verbose=0). Similar configurations for CatBoostClassifier, with “Logloss” or “MultiClass” as loss function.
10f. Internal validation method	• Data were split into training (80%) and testing (20%) sets using train_test_split with random_state=42. Five-fold cross-validation (KFold, shuffle=True, random_state=42) was performed on the training set.
10g. Performance metrics	• Regression (DFS, OS): mean absolute error (MAE) • Classification (DEATH CAUSE): F_1-score for class 1 (tumor-related deaths), prioritizing recall
11. Assessment of prediction performance	• Performance was assessed on the independent test set. For classification, a confusion matrix was used.
12. Model interpretation methods	• CatBoost's built-in feature importance was used. Shapley additive explanations (SHAP) values were computed and visualized using violin plots and scatterplots to understand the individual contributions of each feature.

Results

Patient Characteristics

The study included a final cohort of 374 patients. After excluding incomplete records, the analytical sample sizes were 346 for DFS prediction, 347 for OS prediction, and 312 for death cause prediction. Some records indicate fewer than 374 patients, as not all characteristics were available for every individual in the population.

Table 3 presents the baseline clinical, pathological, and demographic characteristics of the study population, comprising 79.4% (297/374) men and 20.6% (77/374) women. A majority, 79.4% (296/373) were active smokers, and 21.4% (80/374) had a diagnosis of diabetes mellitus. Prior surgery was reported for 33.7% (126/374) of patients, while previous radiotherapy and systemic chemotherapy were less common, 5.1% (19/374) and 3.5% (13/374), respectively. Preoperative hydronephrosis was present in approximately one-third of cases, most frequently unilaterally.

Table . Characteristics of the patients in the dataset.

Characteristic	Value
Continuous variables	
Age (years), mean (SD)	75.2 (9.5)
BMI (kg/m^2), mean (SD)	26.6 (4.2)
Tumor dimension (cm), median (IQR)	2.2 (1.1 - 2.8)
SII[a], median (IQR)	654.7 (408.0 - 1047.0)
DFS[b] (months), median (IQR)	23.0 (6.0 - 52.8)
OS[c] (months), median (IQR)	29.0 (10.8 - 55.4)
Categorical variables, n/N (%)	
Sex	
Men	297/374 (79.4)
Women	77/374 (20.6)
Smoking status	
No	77/373 (20.6)
Yes	296/373 (79.4)
Diabetes mellitus	
No	294/374 (78.6)
Yes	80/374 (21.4)
Prior surgery	
No	248/374 (66.3)
Yes	126/374 (33.7)
Prior radiotherapy	
No	355/374 (94.9)
Yes	197/374 (5.1)
Prior chemotherapy	
No	361/374 (96.5)
Yes	13/374 (3.5)
Pre-hydronephrosis[d]	
None	257/374 (68.7)
Right	44/374 (11.8)
Left	40/374 (10.7)
Bilateral	33/374 (8.8)
Histological examination (H.E. TURV)	
Localized	138/372 (37.1)
Muscle-invasive	212/372 (57)
Squamous	22/372 (5.9)
Urinary diversion type	
Bricker ileal conduit	278/373 (74.5)
Ureterocutaneostomy	36/373 (9.7)
Vesicoileal pouch	59/373 (15.8)
Lymphovascular invasion	
Absent	158/371 (42.6)

Characteristic	Value
Present	213/371 (57.4)
Clinical tumor stage	
cTa	132/366 (36.1)
cTis	43/366 (11.7)
cT1	77/366 (21)
cT2	74/366 (20.2)
cT3	31/366 (8.5)
cT4	9/366 (2.5)
Tumor classification	
T0	16/342 (4.7)
Ta	18/342 (5.3)
Tis	60/342 (17.5)
T1	47/342 (13.7)
T2a	54/342 (15.8)
T2b	7/342 (2)
T3a	80/342 (23.4)
T3b	10/342 (2.9)
T4a	42/342 (12.3)
T4b	8/342 (2.4)
Bladder tumor position	
Intertrigonal zone	44/365 (12.1)
Right periosteal	25/365 (6.8)
Left periosteal	39/365 (10.7)
Dome	22/365 (6)
Posterior wall	58/365 (15.9)
Right lateral wall	71/365 (19.5)
Left lateral wall	58/365 (15.9)
Prostatic urethra	8/365 (2.2)
Anterior wall	26/365 (7.1)
Entire bladder	12/365 (3.3)
Bladder base	2/365 (0.5)
Cause of death	
Alive	205/363 (56.5)
Other	71/363 (19.6)
Cancer	87/363 (23.9)

[a]SII: systemic immune-inflammation index.

[b]DFS: disease-free survival.

[c]OS: overall survival.

[d]Pre-hydronephrosis: pretreatment hydronephrosis.

Histologically, 57% (212/372) of tumors were muscle invasive, while 37.1% (138/372) were localized to the mucosa or submucosa, and 5.9% (22/372) exhibited squamous features. The most common urinary diversion method was Bricker ileal conduit (278/373, 74.5%), followed by vesicoileal pouch construction (59/373, 15.8%) and ureterocutaneostomy (36/373, 9.7%). Lymphovascular invasion was observed in 57.4% (213/371) of patients.

In terms of staging, the most frequent clinical tumor stages were cTa (132/366, 36.1%) and cT1 (77/366, 21%), while advanced stages (cT3 and cT4) were less common (40/366, 11%). Tumor classification was heterogeneous, with T3a (80/342, 23.4%) and Tis (60/342, 17.5%) being most prevalent.

Regarding tumor location, the most frequent sites were the right lateral wall (71/365, 19.5%), the posterior wall (58/365, 15.9%), and the left lateral wall (58/365, 15.9%). At the time of data collection, of 363 patients, 205 (56.5%) were alive, 87 (24%) had died due to cancer-related causes, and 71 (19.6%) had died from other causes.

DFS Prediction

The CatBoostRegressor model was trained to predict DFS in months. Input variables are indicated in Table 1. After 5-fold cross-validation and manual hyperparameter tuning, the model achieved an MAE of 18.68 months, indicating that, on average, the model's predictions deviated by approximately 1.5 years from the observed DFS.

Figure 1 presents the global feature importance ranking from the CatBoost model trained to predict DFS. This ranking reflects the contribution of each variable to reducing the model's prediction across all patients. The most influential predictor was the clinical tumor stage, with an importance score of approximately 17, followed by the pathological tumor classification, with an importance score of approximately 14, reflecting the role of tumor invasiveness, local extension, and accurate tumor staging in DFS prediction. The systemic immune-inflammation index (SII) ranked highly, with a score of approximately 9.5. To a lesser extent, demographic and anatomical variables, such as BMI, age, and tumor dimension, also contributed to the model.

Figure 1. Global feature importance ranking to predict disease-free survival. CTS: clinical tumor stage; DM: diabetes mellitus; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.

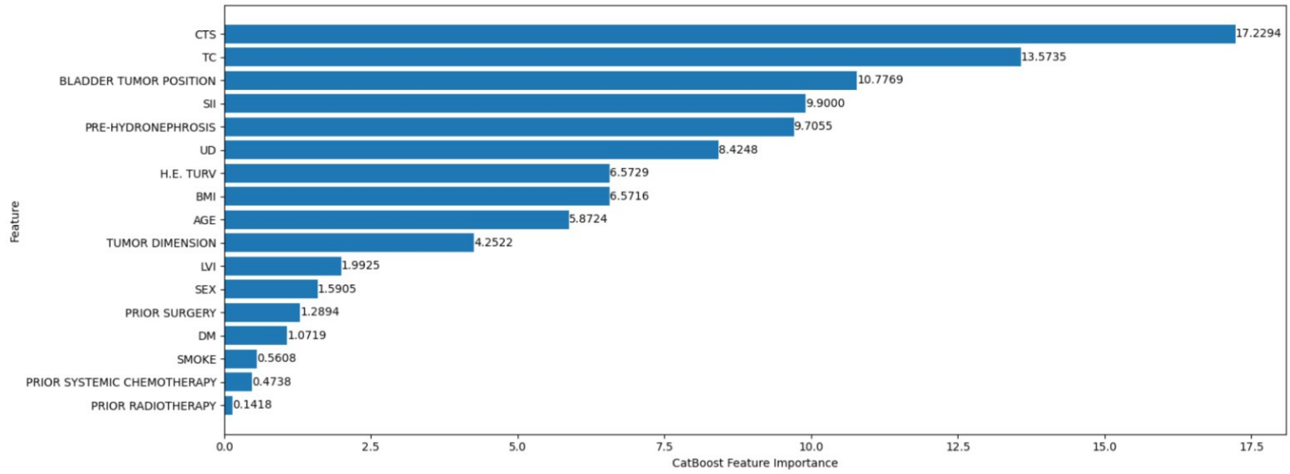


Figure 2 displays the SHAP summary plot for the DFS model, illustrating the distribution and direction of impact of each feature on the predicted DFS across all patients. Clinical tumor stage and tumor classification exhibited the widest distribution of SHAP values, confirming their dominant influence, where the predicted DFS substantially increased or decreased depending on their values. SII displayed a more balanced distribution, with both positive and negative effects depending on the value. In contrast, features such as prior treatment (eg, surgery, radiotherapy, and chemotherapy) and lifestyle factors (eg, smoking status and diabetes) had a SHAP distribution clustered near 0, indicating limited predictive power.

Figure 2. Violin plot of feature influence on disease-free survival prediction from the Shapley additive explanations (SHAP) analysis. CTS: clinical tumor stage; DM: diabetes mellitus; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.

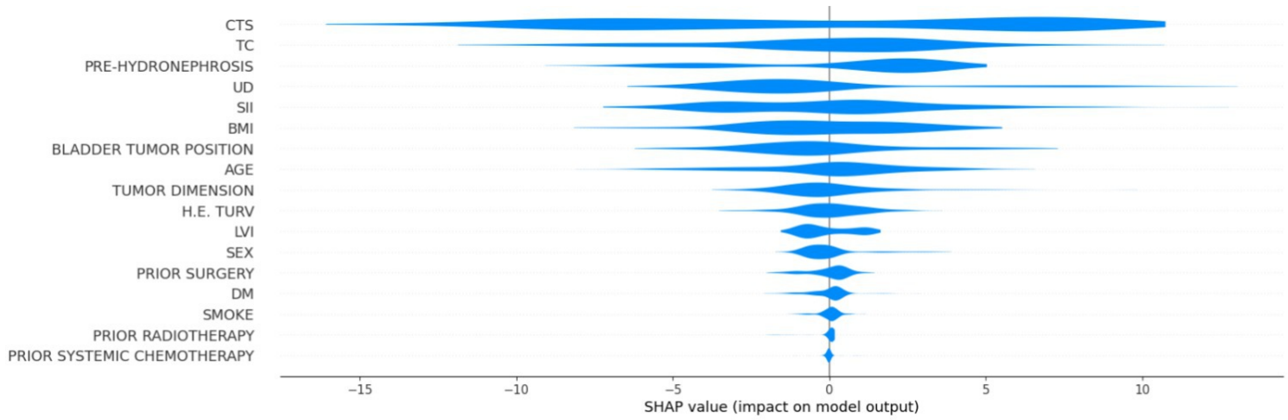
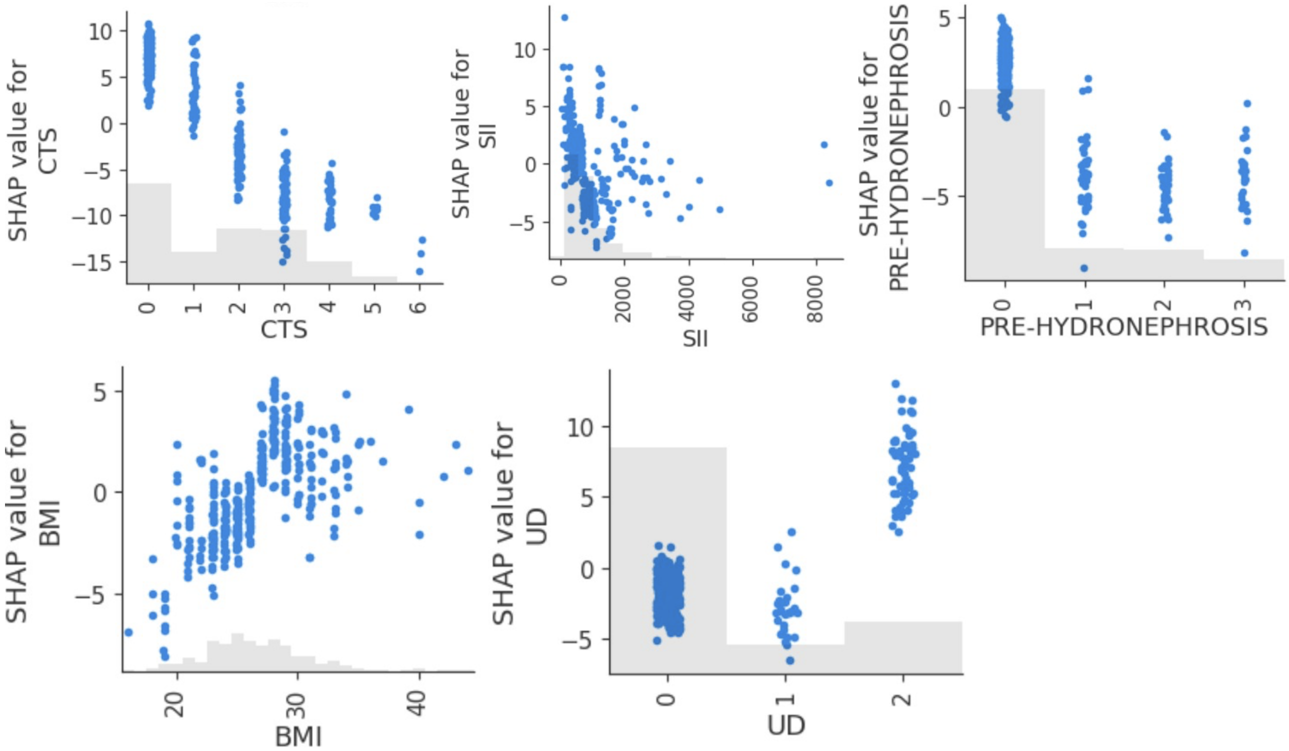


Figure 3 presents the SHAP dependence plots for 4 of the most influential features affecting DFS predictions. The x-axis shows the feature value, and the y-axis shows the SHAP value (ie, the impact on the model's output). Clinical tumor stage showed a strong negative relationship with predicted DFS: as tumor stage increased, the SHAP values shifted sharply downward, indicating a consistent reduction in predicted DFS, aligning with the known prognostic role of tumor invasiveness in bladder cancer. SII demonstrated a nonlinear relationship, showing that patients with lower SII values had better SHAP values, while those with elevated SII showed increasingly negative impacts on DFS. This suggests a threshold effect, where systemic inflammation beyond a certain level contributes to poorer prognosis. The presence of pretreatment hydronephrosis had a negative impact on DFS prediction. Patients with low BMI had negative SHAP values, indicating reduced DFS, while those with moderate BMI experienced mildly negative predictions. At a BMI greater than 28, the SHAP values became positive, suggesting a potential protective effect exerted by higher BMI. Regarding the type of urinary diversion, vesicoileal pouch construction showed positive SHAP values, while other approaches had negative SHAP values.

Figure 3. Shapley additive explanations (SHAP) scatterplots for the 5 most significant features influencing disease-free survival predictions (with BMI in kg/m^2). CTS: clinical tumor stage (0: cTa, 1: cTis, 2: cT1, 3: cT2, 4: cT3, 5: cT4); PRE-HYDRONEPHROSIS: pretreatment hydronephrosis (0: no, 1: right hydronephrosis, 2: left hydronephrosis, 3: bilateral hydronephrosis); SII: systemic immune-inflammation index; UD: urinary diversion type (0: Bricker ileal conduit, 1: ureterocutaneostomy, 2: vesicoileal pouch).



OS Prediction

For OS prediction, the CatBoost model achieved an MAE of 17.2 months across the entire patient cohort. When the analysis was restricted to the subgroup of patients who had died (n=156), the prediction error improved to 15.8 months. After filtering features by importance, using a threshold of <0.5, the MAE further improved to 14.6, suggesting that a more compact feature set may improve predictive efficiency without compromising accuracy (Figure 4). This final model was selected for interpretation, as it maintained accuracy while reducing complexity.

Figure 4. Progressive improvement in model accuracy. MAE: mean absolute error.

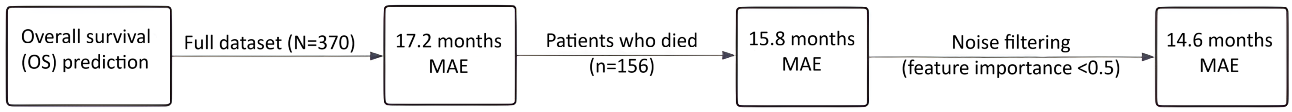


Figure 5 presents the CatBoost feature importance ranking for the best-performing OS prediction model. Tumor classification emerged as the most influential predictor of OS in the final model with a value of approximately 17.5. SII followed closely with a value of approximately 15.5, highlighting the role of systemic inflammation in cancer progression and survival outcomes. The third feature was represented by histological findings (H.E. TURV). Other influencing factors included clinical tumor stage, pretreatment hydronephrosis, type of urinary diversion, BMI, and age.

Figure 5. CatBoost feature importance ranking for the best-performing overall survival prediction model. CTS: clinical tumor stage; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.

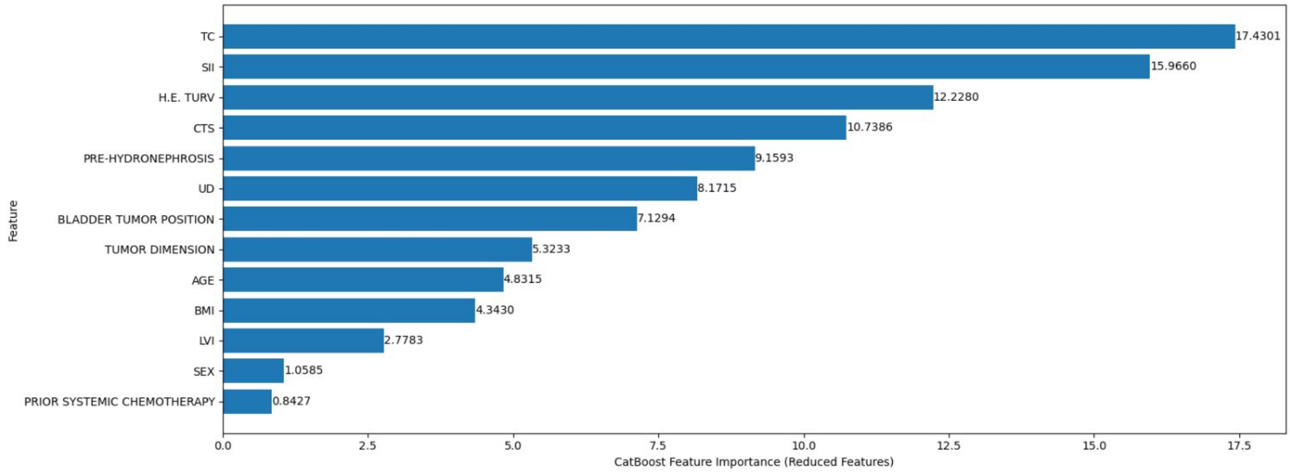


Figure 6 displays the SHAP summary plot for the final OS prediction model. As expected, the most impactful variable was tumor classification, which showed a broad distribution. Pretreatment hydronephrosis and SII exhibited a wide SHAP distribution. Clinical tumor stage and histological findings (H.E. TURV) showed a similar overall effect on prognosis.

Figure 6. Shapley additive explanations (SHAP) violin summary plot for the final overall survival prediction model. CTS: clinical tumor stage; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.

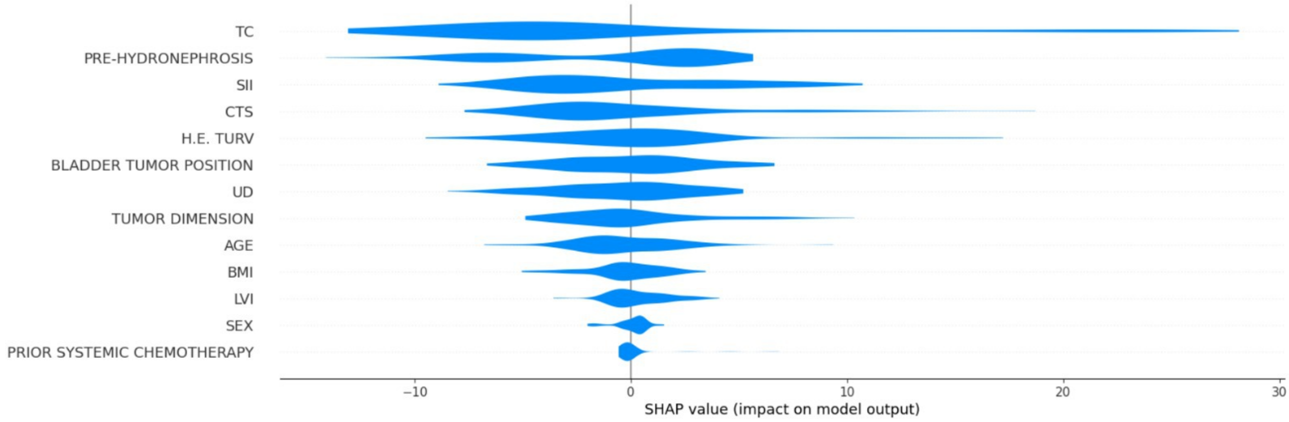
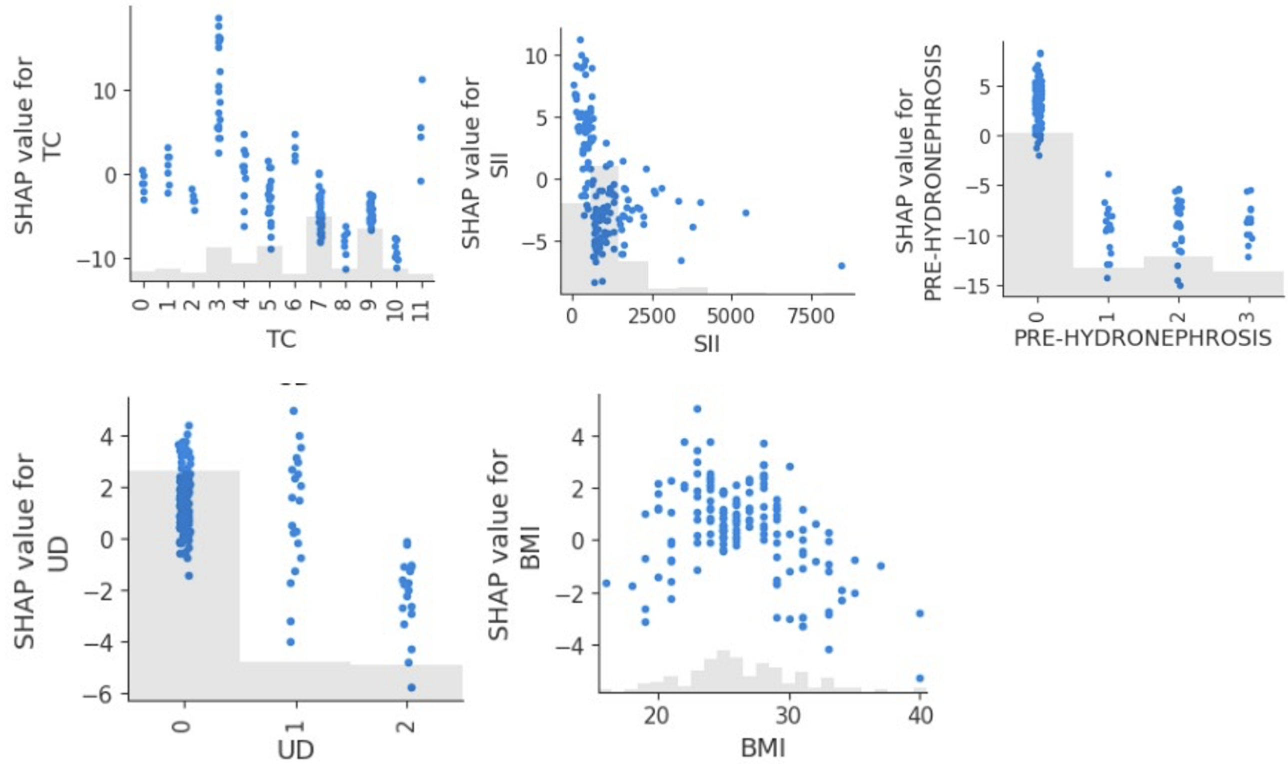


Figure 7 presents the SHAP dependence plots for 5 key features influencing OS predictions. The tumor classification SHAP values indicated that patients with in situ cancers (value 3) had the best OS prediction, which gradually declined as the tumor stage advanced. SII showed a threshold effect, where predictions remained relatively stable up to a value of approximately 1000, then fell sharply, indicating that elevated inflammation is associated with a poor overall outcome. Pretreatment hydronephrosis was strongly linked to reduced predicted OS, where patients with this condition had uniformly negative SHAP values, not influenced by bilaterality. At the same time, BMI demonstrated a nonlinear pattern, where patients with very low BMIs had reduced OS predictions, moderate BMIs were associated with better outcomes, and the SHAP values began to decline again at higher BMI values, suggesting that both underweight and obesity may be associated with increased mortality risk in this population. The type of urinary diversion showed a different impact than that observed in DFS prediction, with vesicoileal pouch construction being associated with a lower OS.

Figure 7. Shapley additive explanations (SHAP) scatterplots for the 5 most influential features influencing overall survival prediction (with BMI in kg/m^2). PRE-HYDRONEPHROSIS: pretreatment hydronephrosis (0: no, 1: right hydronephrosis, 2: left hydronephrosis, 3: bilateral hydronephrosis); SII: systemic immune-inflammation index; TC: tumor classification (1: T0, 2: Ta, 3: Tis, 4: T1, 5: T2a, 6: T2b, 7: T3a, 8: T3b, 9: T4a, 10: T4b); UD: urinary diversion type (0: Bricker ileal conduit, 1: ureterocutaneostomy, 2: vesicoileal pouch).



Cause-of-Death Classification

The CatBoostClassifier was trained to predict whether a patient's death was tumor related. Due to class imbalance, only 14 of 78 deaths were cancer related; custom class weights and a reduced decision threshold of 0.12 were applied to maximize recall and minimize false negatives. The final model achieved a recall of 78.6% (Figure 8), correctly identifying 11 of 14 tumor-related deaths. The overall F_1-score for the positive class was 0.44, with a precision of 31%. The model prioritizes sensitivity over specificity.

Figure 8. Confusion matrix for cause-of-death classification.

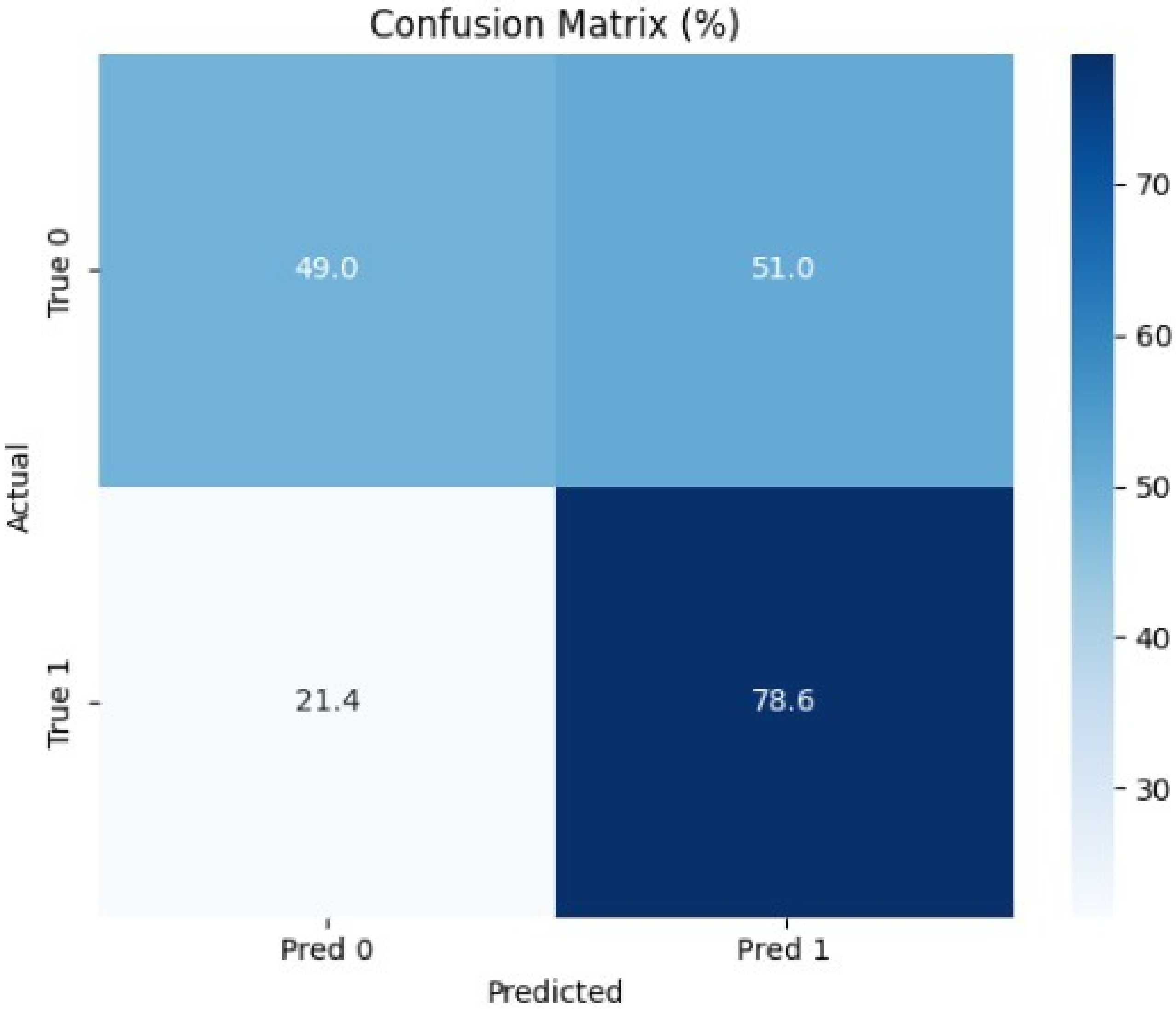
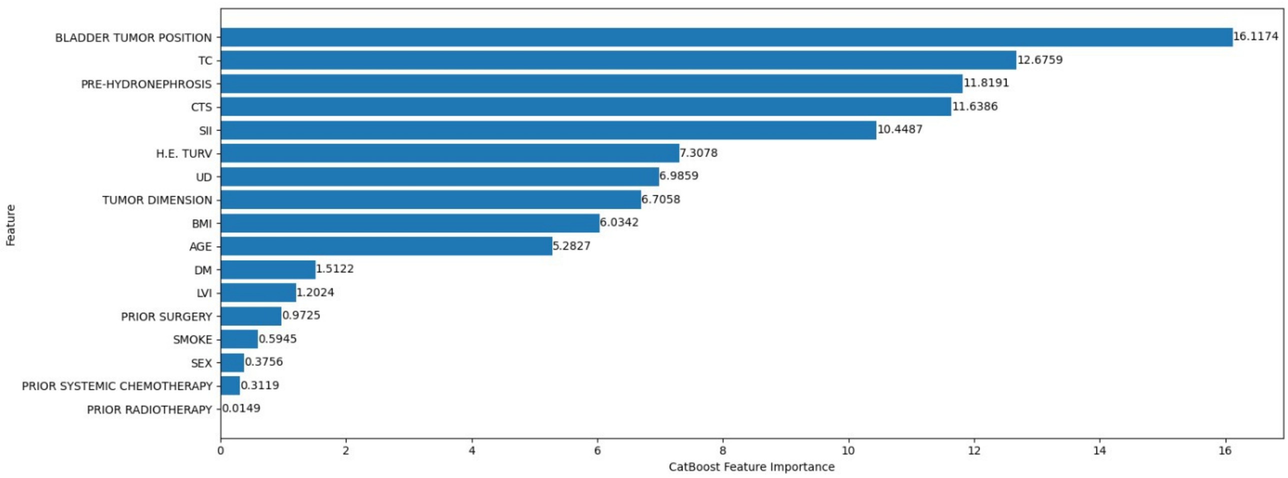


Figure 9 represents the CatBoost feature importance ranking for the death cause classification model. The most influential feature was the anatomical position of the bladder tumor, with a value of approximately 16.5. This was followed by tumor classification and pretreatment hydronephrosis, both indicators of disease severity and progression.

Figure 9. CatBoost feature importance ranking for the cause-of-death classification. CTS: clinical tumor stage; DM: diabetes mellitus; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.



Other key features included the clinical tumor stage and SII, with values of approximately 11 and 11.5, respectively.

Figure 10 displays the SHAP summary plot for the tumor-related death classification model. The feature with the widest and most impactful distribution was bladder tumor position, which is addressed in detail in the discussion of Figure 11. SII was also influential, with positive and negative SHAP values. Tumor classification, pretreatment hydronephrosis, and clinical tumor stage generally acted to slightly decrease the predicted risk of tumor-related death for most patients, with little variability in their effect.

Figure 10. Shapley additive explanations (SHAP) summary plot for the tumor-related death classification model. CTS: clinical tumor stage; DM: diabetes mellitus; H.E. TURV: histological examination for transurethral resection of the bladder; LVI: lymphovascular invasion; PRE-HYDRONEPHROSIS: pretreatment hydronephrosis; SII: systemic immune-inflammation index; TC: tumor classification; UD: urinary diversion type.

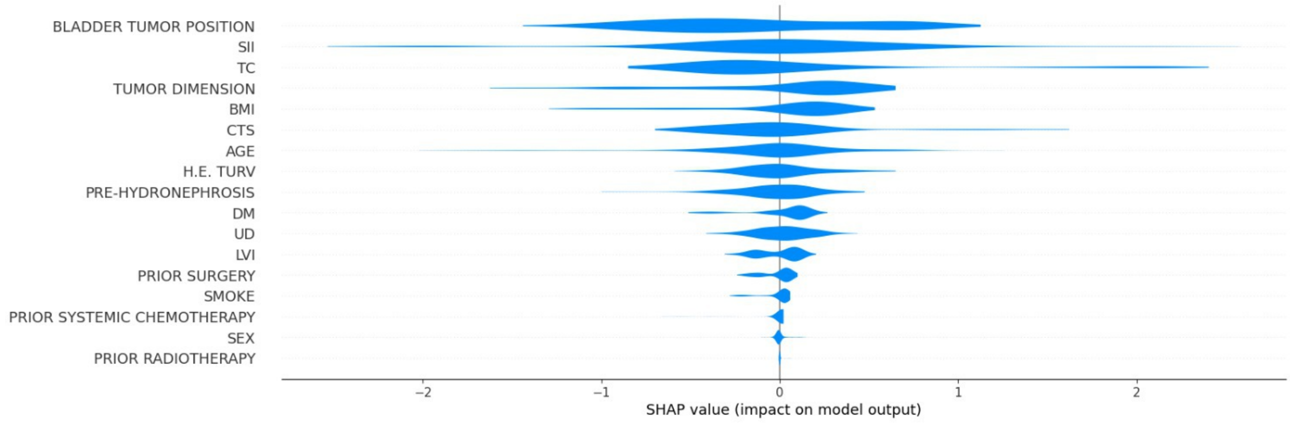


Figure 11. Shapley additive explanations (SHAP) scatterplots for the 5 most influential features influencing the tumor-related death classification model (with BMI in kg/m^2, AGE in years, and TUMOR DIMENSION in cm; bladder tumor position: 0: intertrigonal zone, 1: right periosteal, 2: left periosteal, 3: dome, 4: posterior wall, 5: right lateral wall, 6: left lateral wall, 7: prostatic urethra, 8: anterior wall, 9: entire bladder, 10: bladder base). SII: systemic immune-inflammation index.

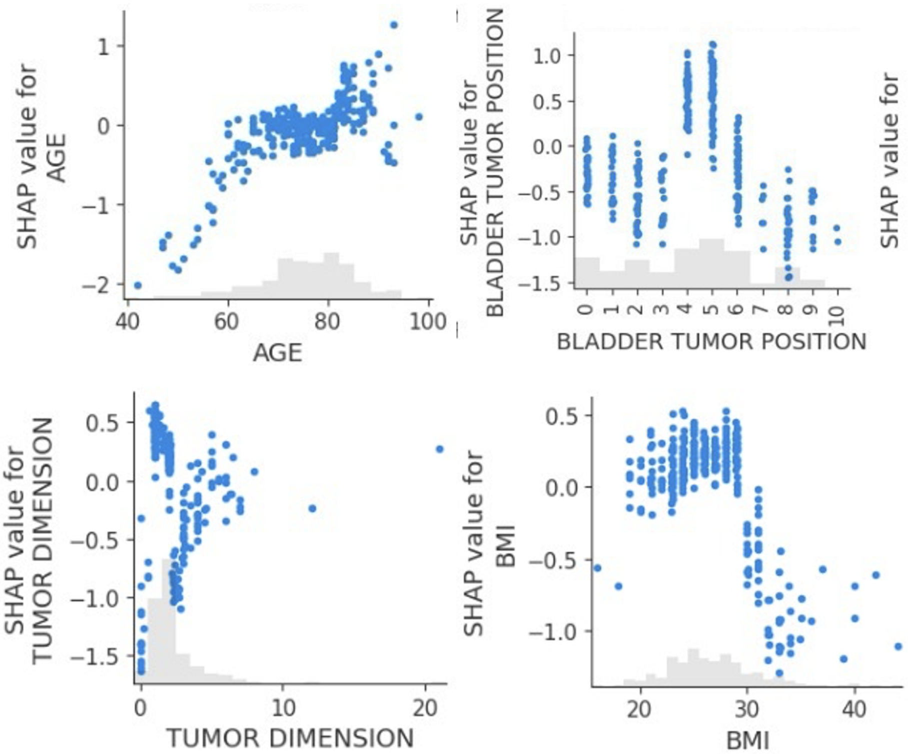


Figure 11 presents the SHAP dependence plots for 5 key features influencing the tumor-related classification model: age, bladder tumor position, SII, tumor dimension, and BMI.

The scatterplot for age exhibited a clear positive trend, showing that as patient age increased, the SHAP value for age also increased, indicating that older age consistently increased the predicted risk of tumor-related death. Bladder tumor position was another influential predictor in the tumor-related death classification model, highlighting that tumors located in the posterior wall and right lateral wall were more likely to be the cause of death. In contrast, tumors located in the anterior wall and at the bladder base or those spreading throughout the entire bladder were less likely to be the cause of death. Increasing SII and tumor dimension also had a moderate predictive value, respectively increasing and decreasing the probability that the patient's cause of death was cancer. Finally, patients with a higher BMI showed a higher likelihood that the cause of death was cancer.

Discussion

Principal Findings

This paper presents the development of a survival prediction model using machine learning approaches for patients with bladder cancer. Our findings demonstrate that modern predictive algorithms show promising accuracy in forecasting DFS, OS, and cause of death. The limited sample size and the paucity of included categories for the analysis suggest that predictive algorithms trained with additional data and variables significantly improve the demonstrated accuracy.

The observation that age showed a positive correlation with survival outcomes is particularly intriguing and seemingly counterintuitive. This "age paradox" has been previously described in bladder cancer and other oncological settings and may reflect a combination of selection bias and underlying tumor biology rather than a true protective effect of age, which can be explained by several factors. Older patients often receive more conservative treatment, which potentially leads to selection bias in surgical candidates. Additionally, younger patients with bladder cancer have been reported to present more frequently with aggressive disease variants, which could account for their relatively poorer outcomes despite younger age [11]. These results align with recent systematic reviews and meta-analyses on radical cystectomy, which consistently report high complication rates and significant variability in postoperative survival across risk profiles [12]. Importantly, as this study is observational, the association between age and survival should be interpreted cautiously and not as a causal relationship.

Clinical tumor stage was the strongest predictor, which aligns with established prognostic factors in bladder cancer [13]. Additionally, inflammatory markers, particularly SII, showed a negative correlation with survival outcomes, supporting recent findings in other malignancies [14]. This relationship likely reflects the complex interplay between systemic inflammation and cancer progression, where elevated SII indicates a protumoral inflammatory state [15]. Our findings on systemic inflammatory indices are consistent with recent data indicating that platelet-to-lymphocyte ratio, systemic inflammation response index, pan-immune-inflammation value, SII, and neutrophil-to-lymphocyte ratio are associated with adverse outcomes in non–muscle-invasive bladder cancer [16].

SHAP analysis revealed a clear, monotonic decline in predicted survival with advancing clinical tumor stage, reinforcing its primary prognostic role in both DFS and OS. SII, by contrast, demonstrated a threshold effect where values above approximately 1000 were associated with a sharp drop in predicted DFS, suggesting a nonlinear relationship between systemic inflammation and patient outcomes. The observed association between urinary diversion type and survival outcomes should be considered exploratory. Previous studies had found that orthotopic neobladder reconstruction had a protective effect against urethral recurrence in male patients undergoing radical cystectomy for bladder cancer [17]. While this effect was not observed in our dataset, we found a positive association between vesicoileal pouch construction and improved survival. This association may reflect both patient selection and potential physiological advantages of this diversion type; however, given the limited number of patients within each diversion subgroup, it should be interpreted cautiously. Confounding factors such as surgical expertise and patient characteristics, which were not accounted for in the present study, may have influenced these findings. BMI showed a similar intriguing relationship: patients with an unhealthy BMI, either high or low, showed poorer outcomes; this may be related both to tumor characteristics and the surgical approach being limited in terms of radicality. This U-shaped association between BMI and survival was consistently observed across both DFS and OS outcomes, with moderate BMI ranges correlated with more favorable SHAP values. The findings support a metabolic vulnerability in patients with underweight as well as obesity, which may influence recovery or treatment tolerance.

Our machine learning models achieved prediction accuracies comparable to those reported in previous studies. The accuracy in cause-of-death prediction, although modest, represents an encouraging level, given the limited resources and the paucity of categories considered for the analysis, particularly when compared with studies published a few years ago that used significantly larger samples yet achieved marginally higher accuracy in mortality and recurrence prediction [18]. A recently published systematic review investigating machine learning algorithms for bladder cancer cystectomy outcomes found that most of the algorithms did not exceed 70% accuracy and, in some cases, performed with approximately 60% accuracy [19]. The integration of SII into predictive models represents an auspicious direction. As a low-cost, readily available biomarker, SII could enhance current prognostic tools without adding significant complexity or cost to patient evaluation [14].

A notable limitation of our study is the relatively high MAE in survival predictions. These MAE values render the algorithm unsuitable for precise individual patient counseling or treatment planning where accurate timing is critical, such as in emergency settings or for patients exhibiting postoperative complications [20]. However, this level of accuracy remains acceptable for clinical trial patient stratification and allocation, particularly in trials where broad risk categories rather than precise survival estimates are needed for randomization. Such applications include balancing treatment arms in clinical trials by identifying comparable risk groups or supporting enrollment decisions in competing risk analysis, where precise timing is less critical than overall risk assessment [21].

Limitations and Reproducibility

This study is subject to some limitations when interpreting the results. The relatively limited dataset size (N=370 initially; reduced to 312 - 347 for specific analyses) inherently constrains the generalizability and robustness of the developed models. While machine learning algorithms such as CatBoost are robust on smaller datasets, their predictive power can be substantially enhanced with larger cohorts.

Secondly, the monocentric nature of the data collection, originating solely from Fondazione Policlinico Universitario Agostino Gemelli IRCCS in Rome, Italy, introduces a potential for selection bias and limits external validity. Patient characteristics, treatment protocols, and population demographics can vary significantly across institutions and geographical regions. The findings from this study may not be directly transferable to other clinical settings without further validation on diverse, external datasets.

Thirdly, while rigorous data cleaning was performed, the inherent human factors associated with retrospective data extraction from medical records cannot be eliminated.

Conclusion

Our study demonstrates the potential utility of machine learning approaches in predicting bladder cancer outcomes following cystectomy. While the achieved accuracy levels are modest, they align with current literature benchmarks and provide a foundation for future development. The identification of clinical tumor stage as the primary predictor, along with the consistent negative correlation of SII with survival outcomes, validates these parameters as valuable prognostic indicators. In particular, the SHAP analysis revealed a monotonic decline in predicted DFS and OS with advancing clinical tumor stage, reaffirming its role in risk stratification. On the other hand, SII exhibited a threshold effect, where values above approximately 1000 were associated with a rapid drop in predicted survival, reinforcing the adverse prognostic impact of systemic inflammation. The current model's performance, though not suitable for precise individual prognostication, shows particular promise for clinical trial stratification and cohort allocation. Future studies with larger datasets and additional predictive variables may enhance the model's accuracy and broaden its clinical applications. Integrating readily available biomarkers, such as SII, represents a cost-effective approach to improving prognostic tools. These findings contribute to the growing body of evidence supporting the role of machine learning in oncological decision-making while acknowledging the need for continued refinement and validation in larger cohorts.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

To ensure the reproducibility of our results and facilitate further research, all code used for this analysis has been made publicly available on Figshare [10].

Authors' Contributions

FAC, BR, and PR elaborated on the first manuscript concept. AN and FAC performed the statistical analysis. FAC, VDV, and PR wrote the article. MS, NF, and GM reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Anagnostou T, Remzi M, Lykourinas M, Djavan B. Artificial neural networks for decision-making in urologic oncology. Eur Urol 2003 Jun;43(6):596-603. [doi: 10.1016/s0302-2838(03)00133-7] [Medline: 12767358]
2. Schwarzer G, Schumacher M. Artificial neural networks for diagnosis and prognosis in prostate cancer. Semin Urol Oncol 2002 May;20(2):89-95. [doi: 10.1053/suro.2002.32492] [Medline: 12012294]
3. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: results from medical data sets. Cancer 2001 Apr 15;91(8 Suppl):1636-1642. [doi: 10.1002/1097-0142(20010415)91:8+<1636::aid-cncr1176>3.0.co;2-d] [Medline: 11309761]
4. Zi H, Liu MY, Luo LS, et al. Global burden of benign prostatic hyperplasia, urinary tract infections, urolithiasis, bladder cancer, kidney cancer, and prostate cancer from 1990 to 2021. Mil Med Res 2024 Sep 18;11(1):64. [doi: 10.1186/s40779-024-00569-w] [Medline: 39294748]
5. Abbod MF, Catto JWF, Linkens DA, Hamdy FC. Application of artificial intelligence to the management of urological cancer. J Urol 2007 Oct;178(4 Pt 1):1150-1156. [doi: 10.1016/j.juro.2007.05.122] [Medline: 17698099]
6. Di Sarno L, Caroselli A, Tonin G, et al. Artificial intelligence in pediatric emergency medicine: applications, challenges, and future perspectives. Biomedicines 2024 May 30;12(6):1220. [doi: 10.3390/biomedicines12061220] [Medline: 38927427]
7. Mazzotta AD, Burti E, Causio FA, et al. Machine learning approaches for the prediction of postoperative major complications in patients undergoing surgery for bowel obstruction. J Pers Med 2024 Oct 8;14(10):1043. [doi: 10.3390/jpm14101043] [Medline: 39452550]
8. Tokuc E, Eksi M, Kayar R, et al. Inflammation indexes and machine-learning algorithm in predicting urethroplasty success. Investig Clin Urol 2024 May;65(3):240-247. [doi: 10.4111/icu.20230302] [Medline: 38714514]

9. Yin M, Joshi M, Meijer RP, et al. Neoadjuvant chemotherapy for muscle-invasive bladder cancer: a systematic review and two-step meta-analysis. Oncologist 2016 Jun;21(6):708-715. [doi: 10.1634/theoncologist.2015-0440] [Medline: 27053504]
10. Bladder cancer CatBoost analysis. Figshare. URL: https://figshare.com/s/ec2419a33ad7ad05f412 [accessed 2026-01-03]
11. Withrow DR, Nicholson BD, Morris EJA, Wong ML, Pilleron S. Age-related differences in cancer relative survival in the United States: a SEER-18 analysis. Int J Cancer 2023 Jun 1;152(11):2283-2291. [doi: 10.1002/ijc.34463] [Medline: 36752633]
12. Sacco E, Campetella M, Marino F, et al. Preoperative risk factors for failure after fixed sling implantation for postprostatectomy stress urinary incontinence (FORESEE): a systematic review and meta-analysis. Eur Urol Open Sci 2025 Jun;77:58-77. [doi: 10.1016/j.euros.2025.05.012] [Medline: 40546391]
13. Kardoust Parizi M, Enikeev D, Glybochko PV, et al. Prognostic value of T1 substaging on oncological outcomes in patients with non-muscle-invasive bladder urothelial carcinoma: a systematic literature review and meta-analysis. World J Urol 2020 Jun;38(6):1437-1449. [doi: 10.1007/s00345-019-02936-y] [Medline: 31493109]
14. Kou J, Huang J, Li J, Wu Z, Ni L. Systemic immune-inflammation index predicts prognosis and responsiveness to immunotherapy in cancer patients: a systematic review and meta-analysis. Clin Exp Med 2023 Nov;23(7):3895-3905. [doi: 10.1007/s10238-023-01035-y] [Medline: 36966477]
15. Wang Q, Zhu D. The prognostic value of systemic immune-inflammation index (SII) in patients after radical operation for carcinoma of stomach in gastric cancer. J Gastrointest Oncol 2019 Oct;10(5):965-978. [doi: 10.21037/jgo.2019.05.03] [Medline: 31602335]
16. Bizzarri FP, Campetella M, Russo P, et al. Prognostic value of PLR, SIRI, PIV, SII, and NLR in non-muscle invasive bladder cancer: can inflammatory factors influence pathogenesis and outcomes? Cancers (Basel) 2025 Jun 28;17(13):2189. [doi: 10.3390/cancers17132189] [Medline: 40647487]
17. Yu J, Lee CU, Chung JH, et al. Impact of urinary diversion type on urethral recurrence following radical cystectomy for bladder cancer: propensity score matched and weighted analyses of retrospective cohort. Int J Surg 2024 Feb 1;110(2):700-708. [doi: 10.1097/JS9.0000000000000904] [Medline: 38000052]
18. Hasnain Z, Mason J, Gill K, et al. Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. PLoS One 2019;14(2):e0210976. [doi: 10.1371/journal.pone.0210976] [Medline: 30785915]
19. Liu YS, Thaliffdeen R, Han S, Park C. Use of machine learning to predict bladder cancer survival outcomes: a systematic literature review. Expert Rev Pharmacoecon Outcomes Res 2023;23(7):761-771. [doi: 10.1080/14737167.2023.2224963] [Medline: 37306511]
20. Qi S, Kumar N, Farrokh M, et al. An effective meaningful way to evaluate survival models. arXiv. Preprint posted online on Jun 1, 2023. [doi: 10.48550/arXiv.2306.01196]
21. Hassan N, Slight R, Morgan G, et al. Road map for clinicians to develop and evaluate AI predictive models to inform clinical decision-making. BMJ Health Care Inform 2023 Aug;30(1):e100784. [doi: 10.1136/bmjhci-2023-100784] [Medline: 37558245]

Abbreviations

AI: artificial intelligence
DFS : disease-free survival
MAE: mean absolute error
OS: overall survival
SHAP: Shapley additive explanations
SII: systemic immune-inflammation index
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by N Rohatgi; submitted 28.Oct.2025; peer-reviewed by E Tokuc, FP Bizzarri; revised version received 09.Jan.2026; accepted 13.Jan.2026; published 19.Feb.2026.

Please cite as:
Causio FA, De Vita V, Nappi A, Sawaya M, Rocco B, Foschi N, Maioriello G, Russo P
Survival Prediction in Patients With Bladder Cancer Undergoing Radical Cystectomy Using a Machine Learning Algorithm: Retrospective Single-Center Study
JMIR Perioper Med 2026;9:e86666
URL: https://periop.jmir.org/2026/1/e86666
doi:10.2196/86666

Enhancing the User Experience of a Perioperative Digital Health Tool for Information Exchange Using a Human-Centered Design Thinking Approach: Qualitative Observational Study

Charlé Steyl[1], MBChB, DCH, DA, MMed; Carljohan Orre[2], MSc, PhD; Greg Foster[3], MSc, PhD; Hanel Duvenage[4], MSc; Michelle S Chew[5,6], MBBS, PhD; Hyla Louise Kluyts[1], MBChB, MMed, DMed

[1]Department of Anaesthesiology, School of Medicine, Sefako Makgatho Health Sciences University, Pretoria, South Africa

[2]Department of Computer Science and Media Technology, Malmö University, Nordenskiöldsgatan 1, Malmö, Sweden

[3]Department of Information Systems, Rhodes University, Makhanda, South Africa

[4]Safe Surgery South Africa, Johannesburg, South Africa

[5]Department of Anaesthesiology and Intensive Care, Linköping University, Linkoping, Sweden

[6]Department of Perioperative Medicine and Intensive Care, Karolinska University Hospital, Stockholm, Sweden

Corresponding Author:
Carljohan Orre, MSc, PhD
Department of Computer Science and Media Technology, Malmö University, Nordenskiöldsgatan 1, Malmö, Sweden

Abstract

Background: Perioperative patient-reported outcomes (PROs) allow patients to share their experiences of surgical procedures with their health care teams using standardized measures. Despite increasing recognition of their value, PROs are not routinely used in clinical practice, partly due to limited evidence of their impact on traditional clinical outcomes and uncertainty among clinicians about their use. Digital health tools offer a promising way to integrate PROs into clinical workflows and enhance patient-clinician interaction, but their success depends on person-centered design to ensure usability and relevance. Safe Surgery South Africa, a nonprofit organization, developed the Perioperative Shared Health Record (PSHR), a secure web-based tool that enables patients to share personal health information and PROs with their anesthetist and surgeon before and after surgery. Initial implementation revealed significant user experience challenges, which contributed to poor uptake.

Objective: This study aimed to explore factors influencing the PSHR user experience in a low- and middle-income country (LMIC) using human-centered design principles.

Methods: This observational qualitative user experience study followed the 5 design thinking stages: empathize, define, ideate, prototype, and test. Semistructured interviews were conducted with postoperative patients from both the public and private health care sectors, including those with and with no prior experience using the PSHR. Thematic analysis followed the 6-phase framework described by Braun and Clarke and was structured using Karagianni's Optimized Honeycomb user experience model. A problem statement was developed, followed by ideation to explore solutions. Paper prototypes were created, refined, and tested through observation, interviews, and validated usability questionnaires.

Results: In the *empathize* stage, 22 interviews were conducted in the private and public health care sectors in South Africa; 7 participants had previous experience using the PSHR. In the *define* stage, participants emphasized the need for connection, feedback, information, and support through their surgical journey. Contrary to expectations, patients were not discouraged by the length of questionnaires if they perceived them as purposeful. In the *ideate* stage, the team considered user expectations and PSHR integration into care processes. In the *prototype* stage, low-fidelity mock-ups were created and refined into paper prototypes. In the *test* stage, testing with 5 participants highlighted the importance of trust, communication, and user-friendly interfaces. Feedback loops and clinician engagement were identified as key motivators for sustained use. The mean usability questionnaire scores indicated excellent usability and high levels of user satisfaction across most domains.

Conclusions: This study is one of the first to apply human-centered design principles to a perioperative digital health tool in an LMIC setting, addressing usability challenges and patient engagement. Key user experience factors influencing patient engagement included communication, feedback, and access to information throughout the surgical journey. Digital health tools such as the PSHR can strengthen communication and support person-centered perioperative care by integrating PROs into clinical workflows and care processes.

(JMIR Perioper Med 2026;9:e79349) doi:10.2196/79349

KEYWORDS

patient-reported outcome measures; person-centered care; patient participation; digital health; perioperative care; user-centered design; user experience research; human-centered design; universal design; design thinking

Introduction

Patients presenting for surgical procedures often feel vulnerable and may become overwhelmed by information that lies outside their usual frame of reference. Many also experience significant physical and emotional symptoms before and after their operation [1,2]. Time constraints and brief interactions during busy ward rounds can limit opportunities for patients to voice concerns or seek clarification [1,3]. In this context, there is a risk that patients feel depersonalized: reduced to passive participants in a system rather than active participants in their own treatment [1]. Furthermore, perioperative clinicians such as surgeons and anesthesiologists often prioritize traditional problem-focused postoperative outcomes such as morbidity and mortality rates, which do not necessarily reflect the outcomes that matter most to individual patients [4-6].

Person-centered care addresses these challenges by recognizing the individual behind the patient: a human being with values, emotions, and goals, and by fostering a partnership that supports patient autonomy and active participation in care decisions [7,8]. Evidence suggests that patients who are empowered and engaged in their own health care may have better outcomes [3,9-11]. Perioperative patient-reported outcomes (PROs) provide one means of achieving this by allowing patients to communicate their experiences of surgical procedures using standardized questionnaires, known as patient-reported outcome and experience measures (PROMs and PREMs) [2,4,12-14]. Various PROs have been defined in the perioperative sphere, including patient satisfaction, quality of recovery (a short-term outcome after surgery), and quality of life (a longer-term outcome after surgery) [14]. Perioperative PROMs and PREMs give individual patients a means to communicate how they are recovering after surgery in such a way that it can be compared between patient groups and procedure types [5,14]. The data can be used to track quality of care over time [6].

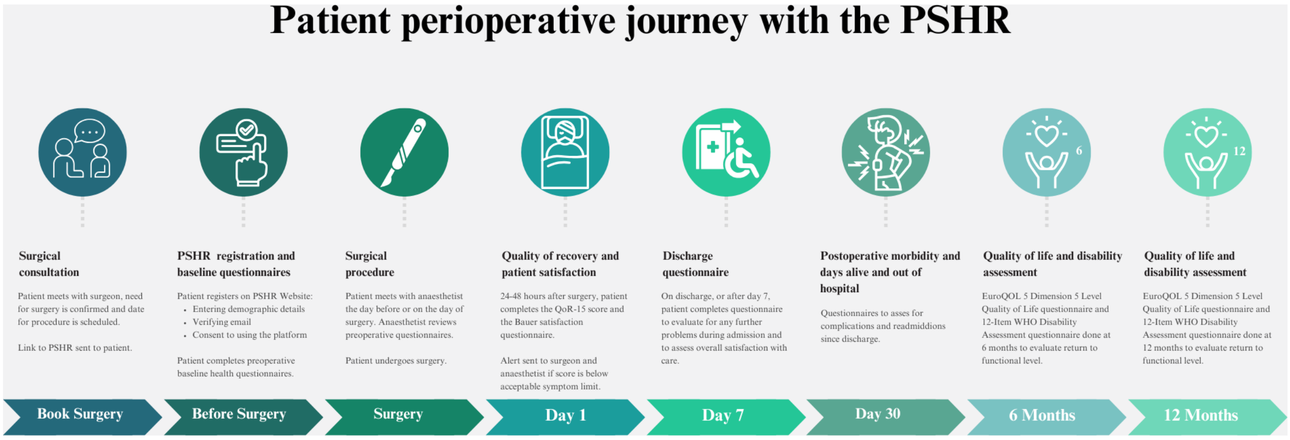
Consensus guidelines recommend the use of PROMs and PREMs in clinical research, and their implementation in perioperative care is increasingly studied [2,6,15-18]. In daily practice, however, the routine use of PROs is hampered by their time-intensive nature, limited evidence linking them to traditional outcomes such as complications or mortality, and uncertainty among clinicians and patients regarding their value [2,6,12,16,17]. These challenges may reflect insufficient person-centeredness in the application of PROs and a lack of responsiveness of health care teams to the information provided by patients [2,19].

Digital health tools offer a promising way to integrate PROs into clinical workflows and to enhance communication between patients and clinicians [17,18,20]. Achieving this, however, requires a human-centered design approach to ensure that digital tools are responsive to diverse user needs and care contexts [21]. Human-centered design forms part of the broader design thinking framework: an empathetic, iterative process that involves end users throughout development to create tools that are understandable, useful, and enjoyable to use [22-25]. Applying these principles through user experience research, which uses interviews, surveys, and usability testing to explore how people interact with digital systems, helps developers create tools that are more intuitive, engaging, and relevant to real-world care [21,26]. Learning from established digital health platforms and implementation of electronic medical record systems can help create digital health tools that support care across the patient journey [27-31].

Perioperative digital health tools have shown promise in high-income countries (HICs) [17,18], but their implementation and use remain underexplored in low- and middle-income countries (LMICs). In South Africa, barriers to large-scale adoption of digital health solutions include limited digital literacy and unequal access to technology and internet across social, economic, and geographic groups [32,33]. These disparities reflect broader inequalities within the health system, where a tax-funded public sector provides care for most of the population but is underresourced, while a private sector funded through medical schemes and out-of-pocket payments serves a minority yet absorbs a large share of resources [34-36]. The public sector continues to rely largely on paper-based documentation, with uneven implementation of systems to capture routine health information and limited electronic record keeping compared with the private sector [37-41]. The private sector is data-rich, with more electronic data systems, but its datasets are typically siloed and not routinely accessible to public governance systems, clinicians, or patients [39,42]. Neither sector currently supports routine or large-scale capture of PROs, limiting opportunities to measure and improve perioperative care from the patient perspective. Addressing these challenges requires context-specific digital solutions that can be designed to strengthen perioperative care in South Africa.

In response to these challenges, Safe Surgery South Africa [43], a research-driven nonprofit organization, developed the Perioperative Shared Health Record (PSHR) [44], a web-based digital tool enabling patients to share baseline preoperative data and postoperative PROs with their surgeon and their anesthetist for up to a year after surgery. Preoperative data can be used in risk stratification and shared decision-making, whereas postoperative PRO data can improve patient care. Data are stored on a secure server but are accessible to both patient and clinician. The system was designed to function across the public and private health care sectors to promote broader accessibility. Figure 1 describes the use of the PSHR in capturing the perioperative journey of a surgical patient.

Figure 1. Patient journey when using the PSHR. EuroQOL: European Quality of Life questionnaire; PSHR: Perioperative Shared Health Record; QoR-15: 15-Item Quality-of-Recovery questionnaire; WHO: World Health Organization.



Initial use of the PSHR in the private health care sector, during the South African Collaborative Surgical Outcomes Study (SACSOS; ClinicalTrials.gov NCT05052021), identified numerous user experience challenges that led to low patient and clinician engagement, which reduced the effectiveness of the PSHR. The registration process was cumbersome, requiring active support for patient users to complete it. In addition, some questionnaires were perceived as lengthy and burdensome, potentially discouraging patients from completing subsequent assessments. As many of the questionnaires are standardized tools designed for specific purposes, content modifications were not always feasible.

The aim of this observational study was to determine the factors that influence the patient user experience of the PSHR as a tool to support perioperative care. The primary objective was to evaluate the user experience of patients in South Africa who had previously used the PSHR during SACSOS. Further objectives were to explore the user needs of patients who have no prior experience with the PSHR and to gain deeper insights into the future design requirements of the PSHR.

To achieve this aim, this study used a human-centered design thinking approach, which required a multidisciplinary team that could combine clinical insight, technical expertise, and practical experience. In this study, anesthesiologists contributed their understanding of perioperative workflows and patient-clinician communication. The information systems, computer science, and media technology researchers applied user experience and design thinking principles to translate patient needs into feasible design solutions. One of the anesthesiologists (CS), with research expertise in PROs, and one of the information systems researchers (CJO), with personal experience as both a patient and a hospital representative, brought perspectives that ensured that the patient remained central throughout the course of the project. The researchers brought together expertise from South Africa and Sweden, combining experience with emerging and advanced digital health systems and perspectives from LMIC and HIC settings. These efforts aim to improve the PSHR's usability and provide insights for more person-centered digital health design.

Methods

Ethical Considerations

This observational qualitative user experience study was approved by the Sefako Makgatho Health Sciences University Research Ethics Committee on November 16, 2023 (SMUREC/M/513/2023:IR), and registered on the National Health Research Database on January 28, 2025 (NHRD: GP_202501_070). Written consent was obtained from all participants and patient privacy and confidentiality was respected by deidentifying patient data and removing any identifiable features from images used in publication. Participation was voluntary with no compensation paid to participants.

Setting

The study took place in South Africa, with postoperative patients and carers recruited via purposeful sampling in both the private (insurance-funded) and public (tax-funded) health care sectors. These 2 sectors are vastly different in South Africa, with a different patient demographic and a significant difference in availability of resources.

Research Team

The research team consisted of 4 female researchers and 2 male researchers. Three of the female researchers (CS, HK, and MC) are practicing anesthesiologists. HD has a research and development background. HK is the founder of Safe Surgery South Africa and focuses on data-driven solutions to improve perioperative risk stratification and surgical outcomes. CS has a clinical and research interest in PROs and previously recruited patients for the PSHR as part of the SACSOS study, maintaining professional relationships with these participants. One of the male researchers, CJO, comes from a digital health and media background and is a patient representative on a hospital management board after surviving cancer. CJO is also involved in an online cancer rehabilitation program spearheading the use of PROMs and PREMs to improve care processes. The other male researcher, GF, comes from a user experience and design science research background.

Methodology

This study was informed by human-centered design principles. The design thinking process was used to structure the study around 5 phases: empathize, define, ideate, prototype, and test [23,25]. In keeping with a person-centered approach, participants informed the project from the outset by sharing their perioperative experiences and needs. These insights guided ideation and design decisions by the investigators, and participants were reengaged during prototype testing to evaluate and improve solutions based on earlier input.

Phase 1: Empathize

The first phase of the project focused on empathizing with PSHR users by interviewing 2 distinct groups. The first group, recruited from the private health care sector, had prior experience using the PSHR before and after surgery through SACSOS (group 1: PSHR experience). The second group included individuals from both the public and private sectors who had undergone surgery but who had no prior exposure to the PSHR (group 2: no PSHR experience). Participants were invited to take part via email, telephone, or by in-person invitation.

For ease of reference when presenting participant quotations, each participant is assigned a letter prefix. "P" denotes participants in group 1, who had used the PSHR before, "U" denotes participants in group 2 from the public sector, and "I" denotes participants in group 2 from the private sector.

Semistructured interviews were conducted between November 2023 and January 2024. Through storytelling, empathy maps and patient journey maps were created. Demographic data were recorded in REDCap (Research Electronic Data Capture) [45] and exported to MS Excel (version 2411; Microsoft Corp). The interviews began by exploring all patients' perioperative experiences. Group 1 participants were then asked about their experiences using the PSHR, while group 2 participants received a brief demonstration of the PSHR before discussing their expectations of a digital tool for perioperative information exchange. The interview guide is included in Multimedia Appendix 1.

Interviews were recorded and transcribed using transcription software (Transcribe—Speech to Text, version 4.20.5; DENIVIP Group LLC) on an iPad dedicated to the project. Transcriptions and audiovisual files were stored in a password-protected online folder. Transcriptions were checked for accuracy by CS and HD and reviewed by all investigators before data analysis. Transcriptions of the initial interviews were systematically coded and analyzed thematically using Nvivo (release 14.23.4(49); Lumiere) by CS and HD. The thematic analysis followed the 6-phase framework described by Braun and Clarke [46]. To explore the main themes related to participants' experiences and expectations of the PSHR, responses were systematically coded and categorized using Karagianni's Optimized Honeycomb model [47]. This model, commonly used in user experience research, structures the analysis of how users interact with a product by breaking down their experiences into 3 primary dimensions: Use, Feel, and Think [47-51]. By applying this framework, we identified patterns in the data and gained deeper insight into the factors influencing user experience of the PSHR.

Phase 2: Define

The information obtained during phase 1 was used to create a problem statement and summary of findings.

Phase 3: Ideate

Insights from the initial interviews and the defining phase informed an ideation phase, during which various solutions were brainstormed by CS, HD, HK, CJO, and GF to enhance future implementation of the PSHR and also taking into account the interoperability with electronic health records.

Phase 4: Prototype

Paper prototypes for the PSHR were created in Balsamiq Wireframes for Desktop (version 4.8.1; Balsamiq Studios LLC). Paper prototypes were refined based on research team group discussions and during user testing.

Phase 5: Test

User testing with the paper prototypes took place in December 2024 with patients and carers who were recruited via email, telephonically, and in person, with the aim to recruit both patients and carers who had used the PSHR before ("expert users") to determine whether insights learned from them during initial interviews had improved their user experience. and patients and carers who had no prior experience of the PSHR ("novice users"), to determine their first time user experience with the system.

As all the expert users would be from the private health care sector, novice users were recruited from the public health care sector. To recruit the expert users, attempts were made to contact all 7 participants from group 1; 3 could not be reached at all, 1 initially agreed but later withdrew, and 3 consented and participated. For the novice users, we intentionally sought individuals with no prior exposure to the platform, including through earlier interview phases, to ensure unbiased, first-time user perspectives. This necessitated recruitment of new participants. Eligibility for user testing included being conversant in English, having basic familiarity with mobile phone and computer use and with the use of the internet. Testing was undertaken by 4 investigators (CJO, CS, GF, and HD), 1 acting as the "computer" to change paper "screens" based on user actions, 1 facilitating the scenario, and 2 observing the interaction; sessions were also audio-recorded for later analysis. For ease of reference, "T" denotes participant responses in the user testing phase.

Participants were asked to complete four tasks during the prototype testing: (1) registering and consenting, (2) completing preoperative baseline questionnaires, (3) finding additional information on the PSHR, and (4) completing postoperative quality of recovery and patient satisfaction questionnaires (Multimedia Appendix 2).

The System Usability Scale (SUS) was used to assess usability after the user prototype testing, as this is a well-established tool that has been found to have good reliability to evaluate the usability of digital systems. The SUS is a 10-statement scale for usability of electronic health applications with good reliability (Cronbach α=0.911) and good face validity [52]. The SUS score ranges from 0 to 100 where higher scores indicate greater usability [53]. A mean SUS score of 68 (SD 12.5) represents the average benchmark for digital health apps [54].

User experience and usability aspects were assessed after prototype testing with the User Experience Questionnaire (UEQ), a well-established tool that evaluates 6 aspects including attractiveness, effectiveness, perspicuity, dependability, stimulation, and novelty with 26 pairs of terms that are scored from 1 to 7 [55]. The questionnaire has good construct validity and good reliability (Cronbach α for 5 of the 6 aspects is above 0.7) [55]. Scoring is done with a downloadable tool, with values ranging between −3 (horribly bad) and +3 (extremely good) [55]. Scores should be evaluated against current benchmarks, freely available for download [56,57].

Prototype user testing was analyzed by 4 investigators (CJO, CS, GF, and HD) who took part in the process using interviews and observation. Thematic analysis of the user tests was done based on research team discussions following the user tests. The SUS and the UEQ were scored in MS Excel according to the guidelines in their reference papers [53-55,57].

Results

Phase 1: Empathize

A total of 22 initial semistructured interviews were conducted as part of empathizing with users. Participant demographics are summarized in (Table 1). All the participants had access to a mobile phone; all but 2 participants in group 2 in the public sector had access to the internet on their mobile phone. Most participants (16/22, 73%) usually used their mobile phones for accessing the internet, whereas 4 out of 22 (18%) participants preferred to use a computer for internet access, and 2 out of 22 (9%) participants did not use the internet at all.

Table . Demographic data for semistructured interviews conducted for phase 1: empathize.

ID[a]	Age (years)	Sex	Language	Race	Prior PSHR[b] experience	Education	Surgery
P1	33	Female	English	Black	Yes	After school qualification	Major abdominal
P2	42	Female	English	White	Yes	After school qualification	Major abdominal
P3	62	Female	Afrikaans	White	Yes	After school qualification	Major abdominal
P4	22	Female	English	Black	Yes	After school qualification	Major abdominal
P5	80	Male	Afrikaans	White	Yes	After school qualification	Major abdominal
P6	72	Male	Afrikaans	White	Yes	After school qualification	Major abdominal
P7	46	Male	Afrikaans	White	Yes	After school qualification	N/A[c]—assisted family member
U1	34	Female	Setswana	Black	No	Secondary school not completed	Major abdominal
U2	22	Male	Northern Sotho	Black	No	Secondary school not completed	Major abdominal
U3	43	Female	Tsonga	Black	No	Secondary school completed	Vascular surgery
U4	65	Male	Setswana	Black	No	Secondary school not completed	Vascular surgery
U5	48	Female	Afrikaans	White	No	Secondary school not completed	Bariatric surgery
U6	29	Female	Northern Sotho	Black	No	After school qualification	Bariatric surgery
U7	42	Female	Afrikaans	Colored[d]	No	After school qualification	Bariatric surgery
U8	39	Female	English	Black	No	After school qualification	Bariatric surgery
U9	42	Male	French	Black	No	Secondary school completed	Vascular surgery
U10	45	Female	Setswana	Black	No	After school qualification	Vascular surgery
I1	64	Female	Afrikaans	White	No	After school qualification	Orthopedic
I2	41	Female	Afrikaans	White	No	After school qualification	Bariatric surgery
I3	54	Female	Afrikaans	White	No	After school qualification	Breast surgery
I4	76	Male	Afrikaans	White	No	After school qualification	Orthopedic
I5	42	Male	Setswana	Black	No	After school qualification	Orthopedic

[a]User ID explanation: “P” denotes participants who had used the PSHR before, “U“ denotes participants from the public sector with no prior experience

of the PSHR, and "I" denotes participants from the private sector with no prior experience of the PSHR.

[b]PSHR: Perioperative Shared Health Record.

[c]N/A: not applicable.

[d]In South Africa, the term "Colored" refers to a distinct cultural and ethnic group with mixed ancestry, recognized as a separate demographic category.

Thematic analysis of the interviews identified 3 main themes: Patient Journey (both groups), PSHR Experience (group 1), and PSHR Expectations (group 2), each with subthemes related to patient engagement and user experience. Detailed findings are provided in Multimedia Appendix 3.

Patient Journey

In understanding the patient perioperative journey (main theme), the following subthemes were identified in both user groups: information-seeking behaviors, emotional response, postoperative difficulties, interaction with health care providers, and advice to other patients. Some participants actively sought more information, either by doing online searches or talking to family members or patients who have been through a similar situation.

> *I was checking [online] how long it's going to be the operation. Okay. Yeah. And how was going to be the pain? How I was cut, a lot of it.* [U1]

> *I go and check like, like the food I have to eat. And then the thing I didn't Google about it is the pills the most. But you know that if you want to go look, you can go, you can go find.* [U3]

> *So I had my sister-in-law, who is a general practitioner, check for the results and then she was the one [that told me].* [P2]

> *And I was also following [on social media], uh, people that will talk about their experience, you know.* [P1]

> *I must say the information from other patients helped a lot, knowing what someone else went through, their experiences, how they felt, what the cost implications were, how they paid it, all of that helped a lot.* [I2]

In both the private and public sectors, there were participants who indicated that they avoided looking for any additional information:

> *I don't really want to Google stuff because you always, there's always stuff. Too much information.* [P3]

> *I think that would've scared me off a little bit more if I knew truly what was to come.* [P4]

> *...because you know when you Google things you don't always get the right information. And it can be very scary.* [I3]

> *So, I give up to an extent that I did not even want to stress myself about the Google information. Because others there are just making some speculations.* [U10]

All the participants experienced some form of emotional turmoil in the time after their diagnosis and before they had surgery, with some describing being in denial, feeling helpless, and isolated:

> *I started like shaking and getting worried. Yes. Like now it's getting worse. And like I took it easy, like okay, fine. I went back to work instead of going to the doctor.* [I5]

> *But now it all became too much. It just felt like you take one step forward and like five steps back… It felt like I was in constant pain and I also felt very helpless.* [P1]

> *It was very... because nobody can come in with you and then you're there alone and then they don't communicate well, doctors all the time, some of them.* [P2]

One participant said an uplifting conversation with her surgeon gave her hope before her surgery and this helped her carry on with her treatment:

> *That answer, that one sentence, and with such conviction, uh, brought back my, um, my hope.* [P3]

Participants in both groups described some physical difficulties in the postoperative period:

> *I think the first two weeks were the hardest really. And the vomiting was much worse at home. Yeah. Uh, the pain also from eating was really terrible.* [P4]

> *But that was also the worst thing that I had the operation, because it was very painful! I didn't expect it would be so painful!* [U7]

Two participants commented that being informed and being able to contact their surgeon made the perioperative journey easier.

> *So being informed. Yeah. Makes you feel more reassured.* [P3]

> *The interesting thing about this surgeon's practice, that I have not come across before, is that he gives you a 24hr whatsapp number that you can use any time of the day if you have problems or questions. There is always someone that responds—that is not something that everyone would do.* [I4]

Participants in both the public and private sectors described their interaction with their health care providers in positive terms, and they valued in-person communication:

> *It made such a difference that [the anaesthetist] were there and [she] could explain to us what was going to happen, it made us feel a lot more secure and calm.* [P6]

> *...what I felt was more these people are taking care of me. I was positive. I could see these other people (points to other patients in ward), they are getting more healthy.* [U2]

Participants offered advice to others that reflected both practical and emotional preparation for surgery. Some emphasized the need to prepare physically by doing breathing exercises and maintaining mobility, while others recognized the emotional impact of surgery on the patient and their families.

> *It’s the emotional side of things that takes quite a toll. But not only on [the patient], but on [the family] too. So I think the emotional strain on both was tough.* [P7]

Several advised future patients to listen to their doctors, follow instructions carefully, and trust the care team. Others highlighted the importance of patience and realistic expectations, especially regarding the time needed for healing:

> *I would tell them that it is very important to listen to what the doctor tells you. To stick to the rules...And I would tell them that they shouldn’t be scared to go through with it.* [I1]

> *I would explain my journey the way it is, then they can come here and get that help because it’s a better help than any other.* [U3]

> *As a patient, I will say first thing first you need to be patient. You need like...healing is a mercy. It won't just happen overnight.* [I5]

PSHR Experience

Group 1 included 6 patients and 1 family member who had used the PSHR. Their user experience, analyzed according to the Optimized Honeycomb model, is summarized in (Table 2), with supporting quotations in Multimedia Appendix 3.

Table . Codes related to experience or expectations of the Perioperative Shared Health Record[a].

	Subthemes[b,c]						
	Use			Feel		Think	
	Findable	Accessible	Usable	Desirable	Credible[d]	Useful	Valuable
PSHR[e] experience: Group 1: PSHR exposed; private sector; 7 interviews conducted with 4 women and 3 men (2 Black and 5 White, aged 18 - 80 years)	• WhatsApp link preferred (9) • Email not used frequently (5) • Desktop used initially (1)	• Mobile phone access preferred (6) • Device limitation (2) • Postoperation no access to glasses (1) • Not comfortable with technology (1)	• Easy to use (4) • Technical problems/bugs (4) • Login process difficult (2) • Importance of feedback (2) • Font size too small (1) • Medical language tricky (1) • Loss of interest over time (1)	• Can complete at home (2)	• No codes	• Personal connection to doctor (10) • Means to feedback from doctor (8) • Patients' ability to express their needs (5) • Benchmarking (4) • Ability to give family access to platform (1)	• Improved care (5) • Patient involvement (1) • Altruism (1)
PSHR expectations: group 2: PSHR unexposed; public sector; 10 Interviews conducted with 7 women and 3 men (7 Black, 2 White, 1 Colored[f], aged 22 - 65 years)	• No codes	• WhatsApp link preferred (12) • Device limitation (4) • Mobile phone access preferred (3) • Email not used frequently (3) • No access to phone postoperatively (2)	• No codes	• No codes	• No codes	• Communication channel (9) • Means to feedback (5) • Benchmarking (4) • Efficiency (3)	• Improved care (2) • Personal connection (2)
PSHR expectations: group 2: PSHR unexposed; private sector; 5 interviews conducted with 3 women and 2 men (1 Black and 4 White, aged 41 - 76 years)	• No codes	• WhatsApp or SMS (2) • Email not easily accessible (2) • Email link useful (1) • No access to phone postoperatively (1)	• No codes	• No codes	• Negative feedback may impact care (1)	• Efficiency (3) • Source of information (3) • Curated list of information (1) • Information about doctor (1) • Communication channel (1) • Link to other patients/support groups (1)	• Altruism (6)

[a]Main theme: PSHR Experience (group 1) and PSHR Expectations (group 2).

[b]Codes from the text were sorted according to the 7 aspects of user experience from the Honeycomb Model (Findable, Accessible, Usable, Desirable,

Credible, Useful, and Valuable) [47]. Code names are included under each aspect, with the code count in parentheses.

[c]Subthemes: Use, Feel, and Think according to Karagiani's Optimized Honeycomb Model [47].

[d]Note that the aspect Credible falls under both Feel and Think.

[e]PSHR: Perioperative Shared Health Record.

[f]In South Africa, the term "Colored" refers to a distinct cultural and ethnic group with mixed ancestry, recognized as a separate demographic category.

Most participants preferred to follow WhatsApp links to find their way to the PSHR on their smartphones:

> *I do prefer that it was easy to use on my phone. So if it can be improved, it must still just be improved. Mainly for, for like a smartphone.* [P1]

> *Yes, because on personal email you're not visiting that often, so it gets lost with all the other stuff. So, I believe your preferred communication is WhatsApp.* [P7,]

One participant commented on the potential difficulty of using a smartphone on the first day postoperatively:

> *It was a bit difficult. Different. If you wear glasses and you don't have glasses on, and you're on morphine. But it wasn't, it wasn't impossible.* [P2]

One of the older participants indicated that they were not comfortable with technology, which is a potential barrier to using a digital tool such as the PSHR:

> *No, I'm not so comfortable with my phone. The internet on there is not something that I usually use.* [P5]

Contrary to expectations, the length of the questionnaires was not perceived in a negative light:

> *It did just go on and on and on. Um, I, I think in my head it was just all part of, just part of the process, preparation and the process you had to do, you know, and making sure that everything is fine.* [P1]

> *It was easy to answer. It doesn't take too long.* [P4.]

However, there was some concern about medical jargon and font size:

> *There's some, um, uh, of the wordings and stuff that I really didn't understand.* [P3]

> *I think the only complaint if I need to complain about improvements will be the size of the font perhaps.* [P7]

Participants valued the PSHR for enhancing their engagement and improving the quality of care they received:

> *...[the doctor] was able to quickly know and come back and improve my care, you know?* [P1]

> *You feel that you were more involved in the planning of your care.* [P2]

Feedback from the surgeon or anesthetist emerged as an important motivator for continued use:

> *[The anaesthetist] had read what was going on. She came and she asked what was going on, and I explained that and she worked around it and talked to the staff.* [P1]

> *If I didn't get feedback, I wouldn't have filled in anymore. I would've done the first one and left it at that.* [P2]

Participants appreciated being able to reflect on their recovery:

> *All of this is quite relevant because it lets you think about your own wellbeing and progress.* [P6]

Interestingly, some participants were also motivated by altruism, expressing a desire to help others:

> *...if my information can help somebody else get through a very difficult situation...then I feel it's worth it.* [P3]

Suggestions for improving the PSHR centered around information sharing and being able to contact patients who had been through a similar procedure:

> *I suppose especially for, for large operations, it might help people to know who the anaesthetist is and have like a name and a, a maybe a photograph of your doctor on the system...And maybe info about postoperative care. Because I mean all these ops have different things and I didn't know I was gonna go to need dietary requirements after the first liver operation...So having that as a portal to kind of find information may be useful.* [P2]

> *I think having someone else who knows, you know, what you've been through would be nice. Yeah. They can give you kind of, like a perspective on what to expect.* [P4]

> *I would prefer to see a video, just a more informal video and then follow up with a verbal conversation just before the operation.* [P7]

PSHR Expectations

Following a brief demonstration of the PSHR, 10 public and 5 private sector postoperative patients with no prior experience of the PSHR (group 2) were interviewed about their expectations of a digital information-sharing tool. Their expectations are summarized in the second part of Table 2, with illustrative quotations provided in Multimedia Appendix 3.

Most participants indicated a preference for accessing the PSHR via a WhatsApp link on their smartphones:

> *I think overall on one's phone is just better, it is more accessible.* [I2]

> *We have emails, but we don't use it so much.* [U2]

> *Anything that is easy for you is easy for me. But really Whatsapp is easiest.* [U10]

Participants indicated that they would value features such as curated information, feedback from their surgeon or anesthetist, and the ability to track their recovery progress.

> *I don't want to get the information by doing a google search. I want information that comes from the doctor themselves, so that I know it is correct.* [I1]

> *You know, if I think back to my work again, clients want to be heard… and now in my setting I am not upset about anything, but it may still be nice to be acknowledged, if I fill something in, it would be nice to get a message or a call to confirm that my responses were seen.* [I3]

> *But when you check, keep on checking on your patient, it's good because if I feel something on me, I have to let you know. Then you'll ask me maybe then to come back at hospital. Then you can check that and sort it out.* [U3]

> *I mean, if they know what your baseline is, what my baseline is, how my life, my, my health is, you know, then they'll know how to proceed. With any procedure for that matter.* [U5]

Participants also indicated that they would be motivated to use a tool such as the PSHR by knowing that they would contribute data that could help others:

> *I would actually do it more for the greater good to contribute to ongoing medical knowledge and learning.* [I1]

> *I think I would still contribute my data if I knew it went for a good cause and if my doctor asked for it.* [I2]

> *If it'll help someone with the same problem that I have, it's important to share it.* [I5]

Potential barriers to using the PSHR are high costs of data and low digital confidence:

> *When I'm at home, I don't see that airtime. Because it's a cost of money. It's very expensive.* [U3]

> *...all the fancy phones, the internet, all that stuff, that's not for me.* [U4]

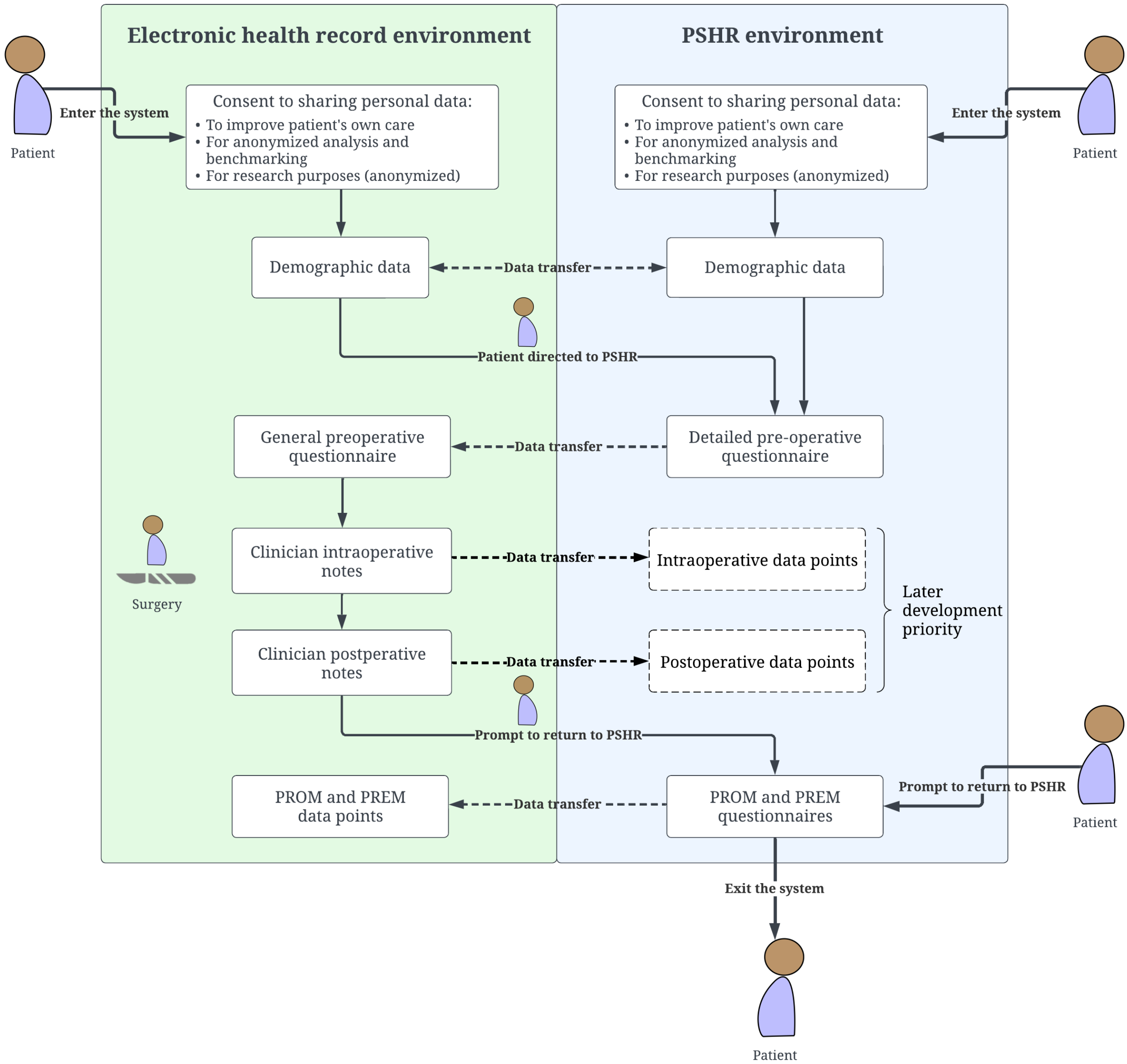
Phase 2: Define

Insights from the initial interviews were that individuals presenting for surgical procedures have a need for connection with and feedback from their health care providers and a willingness to engage in actions necessary to navigate a challenging phase in their lives. Contrary to the investigators' expectations, participants who had used the PSHR were not discouraged by the length of the questionnaires, provided they perceived a clear and meaningful purpose to their completion. In addition, patients expressed a need for information related to their surgical procedures, highlighting the importance of incorporating targeted educational content to support informed decision-making throughout the perioperative journey.

Phase 3: Ideate

The research team brainstormed suggestions and expectations from patient groups, and how the PSHR can be integrated into usual care processes. While standardized questionnaires remained unchanged, their sequence was reorganized to group similar questions, particularly in the PSHR preoperative questionnaire, where multiple risk assessments and surveys are consolidated into a single comprehensive questionnaire. Feedback messages were developed to provide patients with information tailored to their questionnaire responses. Various approaches were explored to support patient-clinician communication through the PSHR. The research team also considered the potential interaction of the PSHR with electronic health records (Figure 2).

Figure 2. Proposed PSHR interaction and interoperability with electronic health records. PREM: patient-reported experience measure; PROM: patient-reported outcomes measure; PSHR: Perioperative Shared Health Record.



Phase 4: Prototype

A series of low-fidelity wireframes were created based on the potential solutions obtained during phase 3. These wireframes were refined into a low-fidelity paper prototype of the PSHR.

Phase 5: User Test

Five individuals consented to participate in prototype testing. Three participants had previously used the PSHR ("expert users"), and 2 participants had no prior experience of the PSHR ("novice users"). Five users have been reported as sufficient for undertaking user testing [58,59]. Demographics are summarized in Table 3. All participants reported having their own mobile phone and usually accessing the internet and their email on their mobile phone and not on a computer. Each testing session took approximately 60 minutes to complete, with the most time spent on the second task. Figure 3 shows paper prototype testing in action; Figure 3A shows a participant discussing task 2 (completing the preoperative questionnaire), and Figure 3B shows a participant responding to a pop-up notification during task 4 (completing postoperative questionnaires).

Table . Demographic data for participants taking part in phase 5: prototype testing.

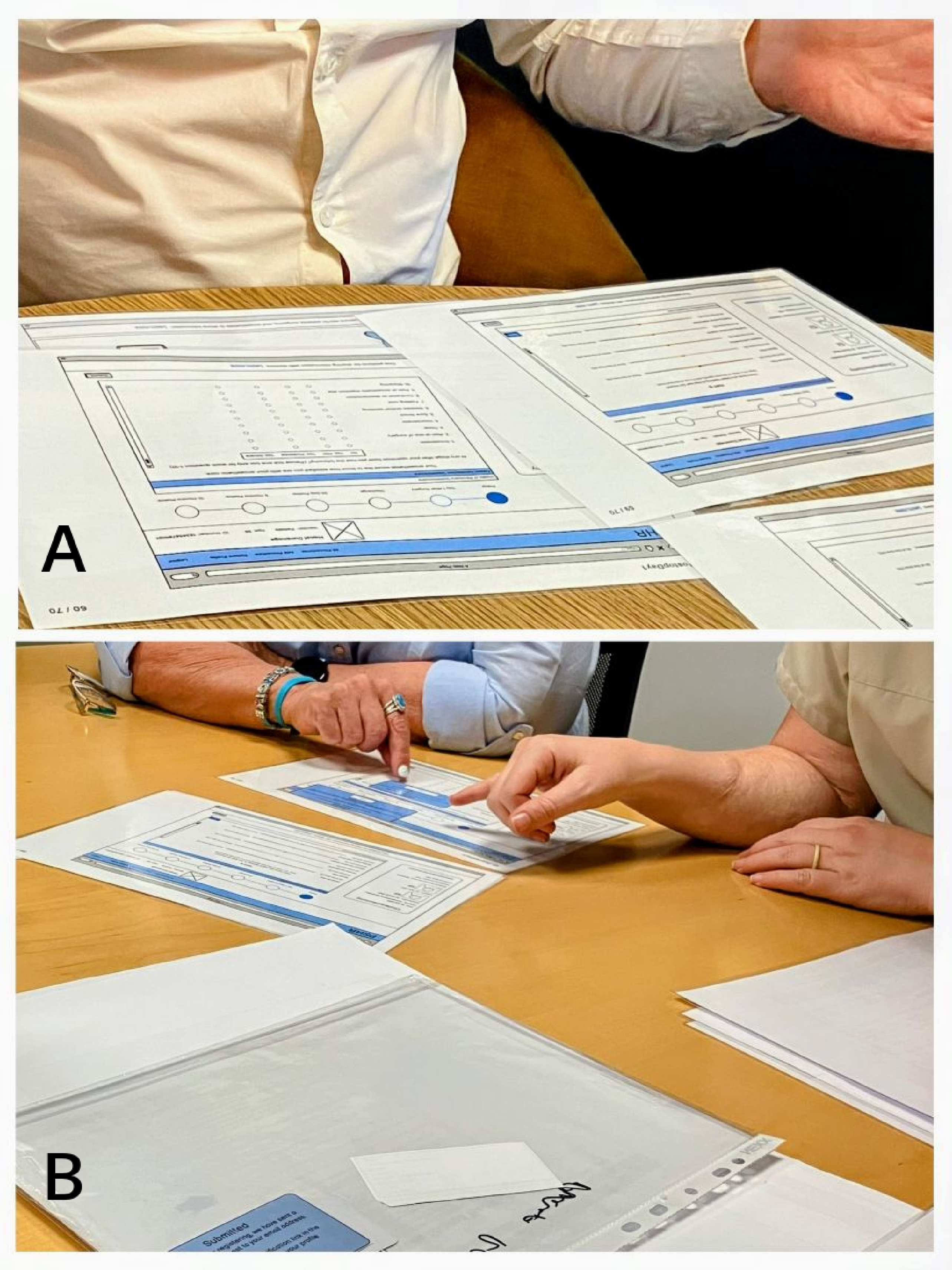
ID[a]	Age (years)	Sex	Language	Race	Prior PSHR[b] experience	Education	Surgery
T1	42	Female	English	White	Yes	After school qualification	Major abdominal
T2	63	Female	Afrikaans	White	Yes	After school qualification	Major abdominal
T3	46	Male	Afrikaans	White	Yes	After school qualification	N/A[c]—assisted family member
T4	21	Male	Northern Sotho	Black	No	After school qualification	Head and neck
T5	32	Male	Setswana	Black	No	Secondary school completed	Head and neck

[a]User ID explanation: “T” denotes participant responses in the user testing phase.

[b]PSHR: Perioperative Shared Health Record.

[c]N/A: not applicable.

Figure 3. Paper prototype testing in action. (A) A participant discussing task 2 (completing the preoperative questionnaire). (B) A participant responding to a pop-up notification during task 4 (completing postoperative questionnaires).



For the first task (registering and consenting), trust was an important factor for 2 of the participants. Prior notification by their doctor to expect a registration email would help improve trust when receiving a link to an unknown website:

> *Yeah, knowing that this is safe, yes, because I spoke to you and I know that you will give me something like this. I think a personal call, direct, to say I'm sending you something now. I trust it more, because I know that I'm protected.* [T4]

Two of the participants read the consent form in detail, 2 participants scrolled through with minimal reading, and 1 participant indicated that he would abort the process when confronted with a lengthy consent process:

> *Unless if I'm buying a car, I don't read the details.* [T1]

> *As soon as I get this information, when I get to this one, I will say, yoh aha, this is too much. Pause, pause, pause.* [T4]

The 2 participants who read the consent form doubted that most users would engage deeply with the consent process. One participant suggested that using illustrations or icons could clarify abstract concepts. During the second task (preoperative questionnaire), various data input methods were tested, including radio buttons, colored numerical sliders, and free text blocks. Users seemed to appreciate color coding to interpret the numerical sliders. One participant suggested modifications to the order of questions.

Participants were able to navigate to the information portal (task 3); all participants found general information links useful but preferred procedure-specific content, with 3 favoring video links over text, and 1 mentioning that they would refer to written information only if the desired content was not available in the video links.

By the fourth task, participants were familiar with the layout of the home page and the questionnaires. Their understanding of the timeline had improved, but several suggestions were made to enhance its visual clarity. Feedback messages following questionnaire completion elicited mixed responses; some participants expressed concern that alerts about poor recovery outcomes could cause anxiety:

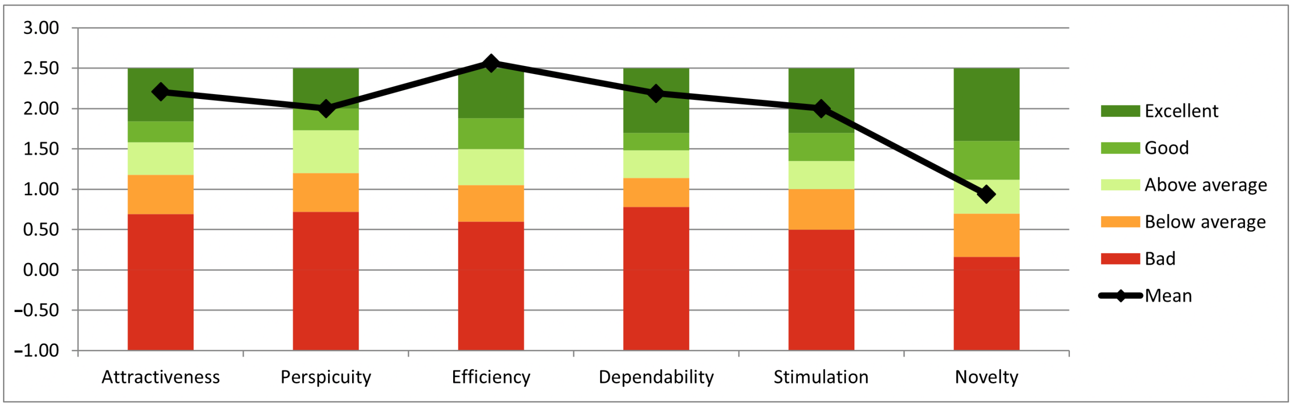
> *It makes me feel worried...I will go back to what I completed [to check] that I completed it correctly. Okay. Because there might be something that I said that might alarm the system.* [T5]

All participants indicated that they would value automatic feedback from their surgeon or anesthetist if they recorded poor scores on their postoperative questionnaires, with expectations for response times ranging from 30 to 60 minutes to up to 48 hours. Participants also noted that a lack of clinician feedback would reduce their motivation to continue using the platform. The potential for the PSHR to enhance patient-doctor relationships is illustrated in this quote:

> *So you still need to get to the buy-in from this. Where's my buy-in coming from? It's coming from the aftercare service, from the doctor, building that relationship. Because [of feedback through the PSHR] I've got a relationship again with the doctor, the surgeon or the rooms. I would be their patient for life! I think that's for me, that's how you get the buy-in to carry on the rest of the process.* [T2]

Four participants (3 experts and 1 novice) completed the SUS and UEQ questionnaires related to the paper prototype testing. The overall mean SUS score was 91.3 (SD 5.7), which indicates very good usability. The mean SUS scores per usability aspect were learnability 91.3 (SD 10.2), efficiency 93.4 (SD 6.9), and satisfaction 89.1 (SD 6.8). The UEQ mean scores for the attractiveness scale are 2.21 (SD 0.6), perspicuity 2.0 (SD 0.9), efficiency 2.6 (SD 0.4), dependability 2.2 (SD 0.6), stimulation 2.0 (SD 0.8), and novelty 0.94 (SD 1.4). The UEQ scores ranged from good to excellent, except for novelty, which scored above average (Figure 4).

Figure 4. Comparing the Perioperative Shared Health Record (PSHR) paper prototype User Experience Questionnaire scores to benchmark data. The measured scale means from our study are compared in relation to existing values from a benchmark dataset, which allows conclusions about the relative quality of the PSHR paper prototypes according to user testing compared with other products.



Discussion

Principal Results

This study applied a human-centered design approach to evaluate and improve patient user experience of a digital health tool developed to capture perioperative patient-reported outcomes. While identifying key usability challenges, it also showed how digital health tools such as the PSHR can help enhance connection between patients and clinicians through information sharing and timely feedback. The findings contribute new evidence from an LMIC setting, where practical integration of PROs into perioperative care is limited. By drawing on patient experiences as a resource for design, the study demonstrates how patient involvement can inform iterative improvements to digital health tools and strengthen person-centered perioperative care [7,21,22,60,61].

The findings of our study align with established user experience principles in digital health design, emphasizing the importance of empathy, communication, accessibility, regulatory compliance, and data privacy and security [21].

Mapping the patient journey revealed the emotional strain of the perioperative period and the value of designing digital health tools that provide empathetic support during this vulnerable time [21,60]. This aligns with previous research highlighting that emotional engagement and good information provision are central to person-centered perioperative care and patient satisfaction [1,62,63]. Experiences reported in HIC show that access to targeted digital health tools can improve patient well-being and empowerment as well as improve postoperative outcomes [18,61,63,64]. Providing patients with a digital resource that offers clear, accessible information may therefore strengthen patient engagement by improving understanding of recovery and fostering a sense of partnership in care [1,18,65].

A key finding was the importance of communication and feedback from clinicians to create trust and to maintain motivation to continue using the system. This aligns with studies on implementation of electronic health records and patient perspectives on digital health tools [61,66]. It was interesting to note that patients were not deterred by lengthy questionnaires if they perceived them as purposeful. Participants valued knowing that their submitted data would inform their care, similar to evidence that perceived purpose and clinician responsiveness may increase adherence to digital platforms [61,64,67]. However, clinicians may not always see the value of using digital health tools to strengthen relationships with patients, especially if these tools are perceived as adding to their workload [28,66,68]. It is important to note that for patients who used the PSHR, a lack of clinician feedback following the completion of the postoperative questionnaires reduced their motivation to continue engaging with the tool. Therefore, it is important that patient needs are balanced with clinical feasibility. One potential solution would be to use automated alerts from PROM data that notify clinicians when patient responses are below a predefined threshold, prompting timely feedback to those patients who need them. This could enhance the perceived usefulness and reliability of the PSHR and strengthen the patient-clinician relationship, without overburdening the clinician [1].

Using the PSHR to track and benchmark recovery progress can offer reassurance or prompt patients to seek help when needed. Such features can promote self-management by helping patients to understand their recovery and to feel more in control of their health [61,64,67]. From our user testing, it emerged that when automated feedback messages to patients flag potential concerns, these messages should balance the communication of information with reassurance and clear guidance on next steps.

Despite the benefits of using a digital tool such as the PSHR, barriers such as limited digital literacy, the high cost of data, and inconsistent access to internet connectivity are significant obstacles to digital health implementation in South Africa [33,39] Patient preference for accessing the PSHR via WhatsApp links highlights the importance of incorporating widely used, low-barrier communication channels in LMIC settings. This aligns with priorities outlined in the South African Digital Health Strategy and the WHO Global Strategy on Digital Health, which highlights the need for equitable access, user-centered design, and interoperability of digital health tools [20,69].

Regulatory constraints, particularly around the consent process, present a design challenge for the PSHR. While simplifying consent forms with icons and condensed text may improve accessibility, maintaining careful attention to detail and to legal and ethical standards is important to ensure the integrity of the consent process. Providing reliable and up-to-date medical information is resource-intensive, whether creating original content or vetting existing material. One potential solution is to involve patients in content development and curation, fostering a collaborative platform. However, ensuring the accuracy of medical content would still require professional oversight and quality control.

An additional consideration in the design of the PSHR is safeguarding data privacy and security, especially as it collects personal and health information subject to the South African Protection of Personal Information Act, comparable with the Health Insurance Portability and Accountability Act in the United States and the General Data Protection Regulation in Europe. In the South African private health care sector, access to patient data can be restricted to the patient and their designated surgeon and anesthetist, which enhances data security. However, in the public health care sector, where care is provided by teams rather than individuals, maintaining data privacy may be more difficult. Concerns about cybersecurity and a lack of trust in an unknown system were seen by some patients as barriers to engaging with the PSHR.

Strengths and Limitations

A key strength of this study lies in its adherence to design thinking and human-centered design principles [21,22]. The use of Karagianni's Optimized Honeycomb model provided a structured lens for analyzing user experience and expectations, capturing functional cognitive and emotional factors influencing user experience [47]. Involving individual patients in the design process enabled the research team to draw on their lived experience as a form of expertise. By prioritizing the needs of actual patients rather than relying on personas, we aim to advance our mission of developing an intuitive, efficient, user-friendly, and person-centered tool. Furthermore, the inclusion of a diverse sample of patients from both the public and private health care sectors in South Africa enhances the study's relevance, particularly as the country progresses toward universal health coverage. Including in the research team a clinician focused on patient-reported outcomes and a hospital patient representative kept the group focused on a patient-centered approach.

The main limitation of this project is the inclusion of a relatively small patient sample across the various phases, limiting the generalizability of findings. However, from a user experience research perspective, idea saturation in initial interviews suggests sufficient theme coverage, and prototype testing revealed consistent usability issues, aligning with Nielsen and

Landauer's 5-user rule [58,59]. The original study protocol aimed to include patients from Sweden to allow comparisons with a high-resource setting, but logistical and resource constraints prevented this. Future research will focus on expanding data collection and user testing, with a comparative analysis between HICs with well-established digital health platforms and LMICs where digital health systems are still being developed. An additional limitation was that user testing was conducted only with low-fidelity paper prototypes. However, this is a recognized approach within design thinking methodology, enabling iterative development without significant cost investment during the early stages when design elements remain subject to change [23,25,70].

Considering that one of the researchers has a professional interest in PROs and had established rapport with several participants, this may have introduced a subtle positive bias in how participants perceived and articulated the value of the PSHR. Furthermore, as the majority of patients had undergone intermediate or major surgical procedures, their emphasis on the need for information and emotional support may not be generalizable to patients presenting for minor operations.

Another limitation is that the study does not capture the user experience of health care providers. The original project plan included workshops with surgeons and anesthetists; however, time constraints necessitated postponing these activities. Ongoing work by the South African authors includes intentional network weaving to promote data-driven surgery which will include engagement with perioperative clinicians.

Although initially unfamiliar to the anesthesiologist investigators, the qualitative and user experience methodologies provided valuable learning. This collaboration highlighted the importance of such approaches in helping clinicians understand patient needs and to develop intuitive digital health tools.

Future Research

Findings from this study will inform further design iterations of the PSHR, both to optimize its use in individual patient care and to generate future research outputs. Next steps include testing a high-fidelity prototype and evaluating the final product in real care settings, with particular attention to patient experiences over time: for example, how patients respond to repeated questionnaires that may appear similar at different intervals. Data from the PSHR will in time become a resource for organizational development and quality improvement. Ongoing development will require the active involvement of both clinicians and patients to ensure that the tool remains relevant, feasible, and responsive to real-world clinical processes and workflows.

Conclusions

This study is one of the first to apply human-centered design principles to a perioperative digital health tool in an LMIC setting, addressing usability challenges and patient engagement. Key user experience factors influencing patient engagement included communication, feedback, and access to information throughout the surgical journey. Digital health tools such as the PSHR can strengthen communication and support person-centered perioperative care by integrating PROs into clinical workflows and care processes. As health care systems worldwide move toward digital integration, our findings provide valuable insights into the factors to consider when digital health tools are introduced in diverse health care contexts. Future research should focus on integrating digital health tools into clinical workflows and assessing their impact on person-centered outcomes and care delivery, with particular emphasis on involving all relevant stakeholders, both clinicians and patients, to ensure that the tools are contextually appropriate and aligned with real-world processes and workflow needs.

Funding

This work is based on the research supported in part by the National Research Foundation of South Africa (reference no. SASUF23041894320) through a seed grant for collaborative research with the South African Swedish University Forum (SASUF-NRF). The research reported in this manuscript is partially supported by the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Research Capacity Development Initiative from funding received from the South African National Treasury. The content and findings reported/illustrated are the sole deduction, view, and responsibility of the researcher, and do not reflect the official position and sentiments of the SAMRC/SASUF-NRF.

Authors' Contributions

CS participated in study concept and design, data collection, data analysis, data interpretation, writing of the manuscript, and critical revision of manuscript. HD contributed to data collection, data analysis, data interpretation, and critical revision of manuscript. CJO and GF participated in study concept and design, data collection, data analysis, data interpretation, and critical revision of manuscript. MC and HK participated in study concept and design and critical revision of manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient semistructured interviews guide.
[DOCX File, 17 KB - periop_v9i1e79349_app1.docx]

Multimedia Appendix 2
Key tasks during paper prototype testing.
[DOCX File, 21 KB - periop_v9i1e79349_app2.docx]

Multimedia Appendix 3
Detailed thematic analysis.
[DOCX File, 57 KB - periop_v9i1e79349_app3.docx]

References

1. Bavin E, Tobiano G, Gillespie BM. Patient experience of elective general surgery: an integrative review. Collegian 2023 Oct;30(5):676-685. [doi: 10.1016/j.colegn.2023.08.002]
2. Wang XS, Gottumukkala V. Patient-reported outcomes: Is this the missing link in patient-centered perioperative care? Best Pract Res Clin Anaesthesiol 2021 Dec;35(4):565-573. [doi: 10.1016/j.bpa.2020.10.006] [Medline: 34801218]
3. Steenkamp L, Silva SA, Collie S. 8-year review of patient satisfaction surveys: Medical scheme members who score initial hospital stay poorly, more likely to have unplanned readmission. 2024. URL: https://www.discovery.co.za/corporate/health-insights-review-patient-surveys-hospitals [accessed 2024-05-02]
4. Roberts GP, Levy N, Lobo DN. Patient-centric goal-oriented perioperative care. Br J Anaesth 2021 Mar;126(3):559-564. [doi: 10.1016/j.bja.2020.12.004] [Medline: 33419527]
5. Myles PS, Boney O, Botti M, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: patient comfort. Br J Anaesth 2018 Apr;120(4):705-711. [doi: 10.1016/j.bja.2017.12.037] [Medline: 29576111]
6. Billig JI, Sears ED, Travis BN, Waljee JF. Patient-reported outcomes: understanding surgical efficacy and quality from the patient's perspective. Ann Surg Oncol 2020 Jan;27(1):56-64. [doi: 10.1245/s10434-019-07748-3] [Medline: 31489556]
7. Ekman I, Swedberg K, Taft C, et al. Person-centered care--ready for prime time. Eur J Cardiovasc Nurs 2011 Dec;10(4):248-251. [doi: 10.1016/j.ejcnurse.2011.06.008] [Medline: 21764386]
8. Leplege A, Gzil F, Cammelli M, Lefeve C, Pachoud B, Ville I. Person-centredness: conceptual and historical perspectives. Disabil Rehabil 2007;29(20-21):1555-1565. [doi: 10.1080/09638280701618661] [Medline: 17922326]
9. Knutsson B, Kadum B, Eneqvist T, Mukka S, Sayed-Noor AS. Patient satisfaction with care is associated with better outcomes in function and pain 1 year after lumbar spine surgery. J Patient Cent Res Rev 2022;9(1):7-14. [doi: 10.17294/2330-0698.1883] [Medline: 35111878]
10. Çakmak C, Uğurluoğlu Ö. The effects of patient-centered communication on patient engagement, health-related quality of life, service quality perception and patient satisfaction in patients with cancer: a cross-sectional study in Türkiye. Cancer Control 2024;31:10732748241236327. [doi: 10.1177/10732748241236327] [Medline: 38411086]
11. Olsson LE, Karlsson J, Ekman I. The integrated care pathway reduced the number of hospital days by half: a prospective comparative study of patients with acute hip fracture. J Orthop Surg Res 2006 Sep 25;1(1):3. [doi: 10.1186/1749-799X-1-3] [Medline: 17150123]
12. Kingsley C, Patel S. Patient-reported outcome measures and patient-reported experience measures. BJA Educ 2017 Apr;17(4):137-144. [doi: 10.1093/bjaed/mkw060]
13. Weldring T, Smith SMS. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights 2013;6(6):61-68. [doi: 10.4137/HSI.S11093] [Medline: 25114561]
14. Moonesinghe SR, Jackson AIR, Boney O, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine initiative: patient-centred outcomes. Br J Anaesth 2019 Nov;123(5):664-670. [doi: 10.1016/j.bja.2019.07.020] [Medline: 31493848]
15. Myles PS, Grocott MPW, Boney O, Moonesinghe SR, COMPAC-StEP Group. Standardizing end points in perioperative trials: towards a core and extended outcome set. Br J Anaesth 2016 May;116(5):586-589. [doi: 10.1093/bja/aew066] [Medline: 27106961]
16. Sokas C, Hu F, Edelen M, Sisodia R, Pusic A, Cooper Z. A review of PROM implementation in surgical practice. Ann Surg 2022 Jan 1;275(1):85-90. [doi: 10.1097/SLA.0000000000005029] [Medline: 34183512]
17. Tsang C, Lee KS, Richards H, Blazeby JM, Avery KNL. Electronic collection of patient-reported outcomes following discharge after surgery: systematic review. BJS Open 2021 Mar 5;5(2):zraa072. [doi: 10.1093/bjsopen/zraa072] [Medline: 33782708]
18. van der Meij E, Anema JR, Otten RHJ, Huirne JAF, Schaafsma FG. The effect of perioperative E-health interventions on the postoperative course: a systematic review of randomised and non-randomised controlled trials. PLoS One 2016;11(7):e0158612. [doi: 10.1371/journal.pone.0158612] [Medline: 27383239]
19. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. BMJ 2014 Mar 26;348:g2225. [doi: 10.1136/bmj.g2225] [Medline: 24671966]
20. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://iris.who.int/server/api/core/bitstreams/1f4d4a08-b20d-4c36-9148-a59429ac3477/content [accessed 2024-08-11]

21. Familoni BT, Odetunde Babatunde S. User experience (UX) design in medical products: theoretical foundations and development best practices. Eng Sci technol j 2024;5(3):1125-1148. [doi: 10.51594/estj.v5i3.975]
22. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021 Jan 12;33(Suppl 1):37-44. [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
23. Krolikowski KA, Bi M, Baggott CM, Khorzad R, Holl JL, Kruser JM. Design thinking to improve healthcare delivery in the intensive care unit: promise, pitfalls, and lessons learned. J Crit Care 2022 Jun;69:153999. [doi: 10.1016/j.jcrc.2022.153999] [Medline: 35104692]
24. Doorley S, Holcomb S, Klebahn P, Segovia K, Utley J. Design thinking bootleg. In: Elbogen E, editor. Standford Design School 2018. URL: https://dschool.sfo3.digitaloceanspaces.com/documents/dschool_bootleg_deck_2018_final_sm2-6.pdf [accessed 2025-12-12]
25. An introduction to design thinking: process guide. Institute of Design at Stanford. URL: https://web.stanford.edu/~mshanks/MichaelShanks/files/509554.pdf [accessed 2025-03-02]
26. Gilbert RM. Reimagining digital healthcare with a patient-centric approach: the role of user experience (UX) research. Front Digit Health 2022;4(August):899976. [doi: 10.3389/fdgth.2022.899976] [Medline: 36016600]
27. Essén A, Scandurra I, Gerrits R, et al. Patient access to electronic health records: differences across ten countries. Health Policy Technol 2018 Mar;7(1):44-56. [doi: 10.1016/j.hlpt.2017.11.003]
28. Cijvat CD, Cornet R, Hägglund M. Factors influencing development and implementation of patients' access to electronic health records-a comparative study of Sweden and the Netherlands. Front Public Health 2021;9:621210. [doi: 10.3389/fpubh.2021.621210] [Medline: 34169054]
29. Wiljer D, Urowitz S, Apatu E, et al. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. J Med Internet Res 2008 Oct 31;10(4):e34. [doi: 10.2196/jmir.1061] [Medline: 18974036]
30. Hägglund M, Scandurra I. Patients' online access to electronic health records: current status and experiences from the implementation in Sweden. Stud Health Technol Inform 2017;245:723-727. [Medline: 29295193]
31. Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish healthcare quality registers. Int J Epidemiol 2015 Oct;44(suppl_1):i174-i174. [doi: 10.1093/ije/dyv096.249]
32. Bridgwood A. General household survey. J Public Health Med 1999 Dec;21(4):482. [Medline: 11469378]
33. Khumalo MH, Moloi TS. Barriers to digital transformation in Gauteng's municipal health clinics. J Local Gov Res Innov 2025 Aug 22;6:1-12. [doi: 10.4102/jolgri.v6i0.282]
34. Reddy CL, Makasa EM, Biccard B, et al. Surgery as a component of universal healthcare: where is South Africa? S Afr Med J 2019 Aug 28;109(9):624-625. [doi: 10.7196/SAMJ.2019.v109i9.14233] [Medline: 31635583]
35. Coovadia H, Jewkes R, Barron P, Sanders D, McIntyre D. The health and health system of South Africa: historical roots of current public health challenges. Lancet 2009 Sep 5;374(9692):817-834. [doi: 10.1016/S0140-6736(09)60951-X] [Medline: 19709728]
36. Day C, Gray A, Cois A, Ndlovu N, Massyn N, Boerma T. Is South Africa closing the health gaps between districts? Monitoring progress towards universal health service coverage with routine facility data. BMC Health Serv Res 2021 Sep 13;21(Suppl 1):194. [doi: 10.1186/s12913-021-06171-3] [Medline: 34511085]
37. Mayosi BM, Lawn JE, van Niekerk A, et al. Health in South Africa: changes and challenges since 2009. Lancet 2012 Dec 8;380(9858):2029-2043. [doi: 10.1016/S0140-6736(12)61814-5] [Medline: 23201214]
38. Thabethe SS, Mathews VE. Quality of routine health data in DHIS2 in South Africa: Eastern Cape province from 2017–2020. S Afr J Inf Manag 2024;26(1):1903. [doi: 10.4102/sajim.v26i1.1903]
39. Zharima C, Griffiths F, Goudge J. Exploring the barriers and facilitators to implementing electronic health records in a middle-income country: a qualitative study from South Africa. Front Digit Health 2023;5:1207602. [doi: 10.3389/fdgth.2023.1207602] [Medline: 37600481]
40. Kluyts HL. Strengthening surgical systems in LMICs: data-driven approaches. Lancet Global Health 2024 Nov;12(11):e1744-e1745. [doi: 10.1016/S2214-109X(24)00375-9] [Medline: 39245052]
41. Nicol E, Hanmer LA, Mukumbang FC, Basera W, Zitho A, Bradshaw D. Is the routine health information system ready to support the planned national health insurance scheme in South Africa? Health Policy Plan 2021 Jun 1;36(5):639-650. [doi: 10.1093/heapol/czab008] [Medline: 33822055]
42. Matshidze P, Hanmer L. Health information systems in the private health sector. In: South African Health Review: Health Systems Trust; 2007. URL: https://www.hst.org.za/publications/South%20African%20Health%20Reviews/11_SAHR_2006-2007_Section%206.pdf [accessed 2025-12-12]
43. Transforming surgical care in Africa through data-driven innovation. Safe Surgery SA. URL: https://safesurgery.co.za [accessed 2020-10-15]
44. Perioperative shared health record. Safe Surgery SA. URL: https://safesurgery.co.za/pshr/ [accessed 2023-04-08]
45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

47. Karagianni K. Optimizing the UX honeycomb. Medium. 2018. URL: https://uxdesign.cc/optimizing-the-ux-honeycomb-1d10cfb38097 [accessed 2024-10-13]
48. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
49. Im D, Pyo J, Lee H, Jung H, Ock M. Qualitative research in healthcare: data analysis. J Prev Med Public Health 2023 Mar;56(2):100-110. [doi: 10.3961/jpmph.22.471] [Medline: 37055353]
50. Mansson L, Wiklund M, Öhberg F, Danielsson K, Sandlund M. Co-creation with older adults to improve user-experience of a smartphone self-test application to assess balance function. Int J Environ Res Public Health 2020 May 26;17(11):3768. [doi: 10.3390/ijerph17113768] [Medline: 32466484]
51. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
52. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
53. Brooke J. SUS: A “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry, 1st edition: Taylor & Francis; 1996.
54. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
55. Laugwitz B, Held T, Schrepp M. Construction and Evaluation of a User Experience Questionnaire: Springer; 2008, Vol. 5298 LNCS:63-76. [doi: 10.1007/978-3-540-89350-9_6]
56. A fast and reliable questionnaire to measure the User Experience of interactive products. User Experience Questionnaire (UEQ). URL: https://www.ueq-online.org/ [accessed 2025-01-19]
57. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the User Experience Questionnaire (UEQ). Int J Interact Multim Artif Intell 2017;4(4):40. [doi: 10.9781/ijimai.2017.445]
58. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. 1993 Presented at: Conference on Human Factors in Computing Systems—Proceedings Published by ACM; Apr 24-29, 1993 URL: http://portal.acm.org/citation.cfm?doid=169059 [accessed 2025-12-19] [doi: 10.1145/169059.169166]
59. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2025-01-19]
60. Molina-Recio G, Molina-Luque R, Jiménez-García AM, Ventura-Puertos PE, Hernández-Reyes A, Romero-Saldaña M. Proposal for the user-centered design approach for health apps based on successful experiences: integrative review. JMIR Mhealth Uhealth 2020 Apr 22;8(4):e14376. [doi: 10.2196/14376] [Medline: 32319965]
61. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients’ perspectives on digital health tools. PEC Innov 2023;2:100171. [doi: 10.1016/j.pecinn.2023.100171]
62. Jones CH, O’Neill S, McLean KA, Wigmore SJ, Harrison EM. Patient experience and overall satisfaction after emergency abdominal surgery. BMC Surg 2017;17(1):76. [doi: 10.1186/s12893-017-0271-5]
63. Sansalone E, White J, Zucca A, et al. Acceptability of the RecoverEsupport digital health intervention among patients undergoing breast cancer surgery: qualitative study. J Med Internet Res 2025 Oct 17;27:e77567. [doi: 10.2196/77567] [Medline: 41105868]
64. ElKefi S, Asan O. How technology impacts communication between cancer patients and their health care providers: a systematic literature review. Int J Med Inform 2021 May;149:104430. [doi: 10.1016/j.ijmedinf.2021.104430] [Medline: 33684711]
65. Hladkowicz E, Yachnin D, Boland L, et al. Evaluation of a preoperative personalized risk communication tool: a prospective before-and-after study. Can J Anaesth 2020 Dec;67(12):1749-1760. [doi: 10.1007/s12630-020-01809-y] [Medline: 32929659]
66. Hägglund M, Scandurra I. Usability of the Swedish accessible electronic health record: qualitative survey study. JMIR Hum Factors 2022 Jun 23;9(2):e37192. [doi: 10.2196/37192] [Medline: 35737444]
67. Esper SA, Holder-Murray J, Meister KA, et al. A novel digital health platform with health coaches to optimize surgical patients: feasibility study at a large academic health system. JMIR Perioper Med 2024 Apr 4;7(1):e52125. [doi: 10.2196/52125] [Medline: 38573737]
68. Ahmed A, Ho CW, Grant Y, Archer S, Carrington EV. Acceptability of digital health interventions in perioperative care: a systematic review and narrative synthesis of clinician perspectives. BMJ Open 2025 Mar 15;15(3):e086412. [doi: 10.1136/bmjopen-2024-086412] [Medline: 40090692]
69. National digital health strategy for South Africa 2019-2024. National Department of Health. 2019 May. URL: https://knowledgehub.health.gov.za/system/files/elibdownloads/2023-04/national-digital-strategy-for-south-africa-2019-2024-b.pdf [accessed 2025-10-20]
70. Altman M, Huang TTK, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117. [doi: 10.5888/pcd15.180128] [Medline: 30264690]

Abbreviations

HIC: high-income country
LMIC: low- and middle-income country
PREM: patient-reported experience measure
PRO: patient-reported outcome
PROM: patient-reported outcome measure
PSHR: Perioperative Shared Health Record
REDCap: Research Electronic Data Capture
SACSOS: South African Collaborative Surgical Outcomes Study
SUS: System Usability Scale
UEQ: User Experience Questionnaire

Edited by J Shiffermiller; submitted 19.Jun.2025; peer-reviewed by J Ohde, S Nilsson; revised version received 28.Oct.2025; accepted 18.Nov.2025; published 12.Jan.2026.

Please cite as:
Steyl C, Orre C, Foster G, Duvenage H, Chew MS, Kluyts HL
Enhancing the User Experience of a Perioperative Digital Health Tool for Information Exchange Using a Human-Centered Design Thinking Approach: Qualitative Observational Study
JMIR Perioper Med 2026;9:e79349
URL: https://periop.jmir.org/2026/1/e79349
doi:10.2196/79349

Clinical Feasibility and Outcomes of Surgeon-Performed Laparoscopic-Guided Subcostal Transversus Abdominis Plane Block in Laparoscopic Cholecystectomy: Prospective Observational Study

Sarun Mahasupachai, MD; Thawatchai Tullavardhana, MD

Department of Surgery, Faculty of Medicine, Srinakharinwirot University, Ransit - Nakhon nayok Road, Ongkharak, Nakhon nayok, Thailand

Corresponding Author:
Thawatchai Tullavardhana, MD
Department of Surgery, Faculty of Medicine, Srinakharinwirot University, Ransit - Nakhon nayok Road, Ongkharak, Nakhon nayok, Thailand

Abstract

Background: Laparoscopic-guided subcostal transversus abdominis plane (TAP) block has been introduced as a surgeon-performed approach to postoperative analgesia in laparoscopic cholecystectomy (LC), allowing direct visual confirmation of local anesthetic delivery without ultrasound guidance. However, evidence regarding its clinical outcomes, particularly in patients with complicated gallstone disease, remains limited.

Objective: This study aimed to evaluate postoperative analgesic outcomes and identify factors associated with opioid requirement following laparoscopic-guided subcostal TAP block.

Methods: A prospective observational study was conducted between November 2023 and October 2024 at Srinakharinwirot University Hospital, Thailand. Patients (aged 18 - 80 years) undergoing LC for uncomplicated or complicated gallstone disease received a laparoscopic-guided subcostal TAP block with 0.25% bupivacaine. Postoperative pain was assessed using the Visual Analog Scale at 2, 4, 6, 8, 12, and 24 hours. Morphine administration within the first 24 hours was recorded. Associations between perioperative variables and opioid requirement were analyzed using univariate and exploratory multivariable logistic regression.

Results: A total of 42 patients were included in the analysis. Of these, 21 (50%) did not require postoperative opioids, while the remaining patients (n=21, 50%) received a mean cumulative morphine dose of 3.86 (SD 1.39) mg within 24 hours. Pain scores were lower during the early postoperative period (2, 4, and 12 h) in patients who did not require opioids. Higher American Society of Anesthesiologists classification was independently associated with postoperative morphine requirement (odds ratio 6.51, 95% CI 1.37 - 30.96; *P*=.01). No major complications or local anesthetic toxicity were observed.

Conclusions: In this prospective observational cohort, laparoscopic-guided subcostal TAP block may be associated with favorable early postoperative analgesic profiles and relatively low opioid requirements after LC, including in patients with gallstone-related complications. Higher American Society of Anesthesiologists classification may be associated with increased opioid demand, highlighting the importance of individualized, risk-adapted analgesic strategies. Although limited by the absence of a control group and modest sample size, these findings support the clinical feasibility of surgeon-performed TAP block for consideration within multimodal analgesia approaches in enhanced recovery after surgery–oriented perioperative care.

Trial Registration: Thai Clinical Trials Registry TCTR20250314002; https://www.thaiclinicaltrials.org/show/TCTR20250314002

(JMIR Perioper Med 2026;9:e87622) doi:10.2196/87622

KEYWORDS

laparoscopic cholecystectomy; minimally invasive surgery; postoperative pain; transversus abdominis plane block; enhanced recovery after surgery

Introduction

Laparoscopic cholecystectomy (LC) is the standard surgical approach for gallstone disease, offering distinct advantages over open cholecystectomy, including shorter recovery times and earlier return to normal activities [1]. Despite these benefits, LC is associated with moderate to severe postoperative pain, particularly within the first 24 hours, often necessitating opioid analgesia [2]. High-dose opioid use, however, is frequently complicated by nausea, vomiting, dizziness, abdominal distension, and urinary retention, which may delay recovery and prolong hospitalization [3,4].

Enhanced recovery after surgery (ERAS) protocols have been widely implemented to optimize perioperative care and expedite recovery. Multimodal analgesia represents a cornerstone of these protocols, aiming to minimize opioid use while maintaining effective pain control [5,6]. Within this framework, the subcostal transversus abdominis plane (TAP) block has emerged as a valuable component of multimodal analgesia, providing targeted pain relief following LC. When performed under ultrasound guidance using 0.25% bupivacaine, this block reliably anesthetizes thoracic (T7-T12) and lumbar (L1) nerves, thereby improving pain control and facilitating earlier mobilization [7,8]. Nevertheless, its dependence on anesthesiologist expertise and specialized equipment limits feasibility in certain clinical environments.

To address these limitations, the laparoscopic-guided subcostal TAP block has been developed as a surgeon-performed technique seamlessly incorporated into the operative workflow. Under direct laparoscopic visualization, local anesthetic can be precisely delivered into the TAP, providing consistent parietal analgesia while obviating the need for ultrasound equipment or additional personnel. This method has been demonstrated to be safe, efficient, and time-effective, offering a practical alternative for postoperative pain control following LC [9].

Nevertheless, most previous studies evaluating TAP block for LC have primarily focused on patients with uncomplicated gallstone disease and were conducted in controlled trial settings. A recent systematic review and meta-analysis demonstrated that TAP block is effective in reducing postoperative pain and opioid consumption after LC, with most included studies using ultrasound-guided techniques [10]. In contrast, this study evaluates a surgeon-performed, laparoscopic-guided subcostal TAP block integrated into routine operative workflow.

This study aimed to assess the clinical feasibility and outcomes of surgeon-performed, laparoscopic-guided subcostal TAP block for postoperative pain management in patients undergoing LC for both uncomplicated and complicated gallstone disease, including acute cholecystitis and biliary tract obstruction. Additionally, perioperative predictors of postoperative opioid requirement were explored.

Methods

Overview

A single-center observational study was conducted at the Department of Surgery, Faculty of Medicine, Srinakharinwirot University, Thailand, between November 2023 and October 2024. Eligible patients were aged 18 to 80 years and diagnosed with symptomatic cholelithiasis or gallstone-related complications. Uncomplicated gallstone disease was defined as symptomatic cholelithiasis or chronic cholecystitis without evidence of systemic inflammation or biliary complications. Complicated gallstone disease was defined as gallstone-related conditions associated with acute inflammation or biliary obstruction, including acute cholecystitis, acute cholangitis, or biliary obstruction requiring endoscopic retrograde cholangiopancreatography. Patients who were converted to open surgery or had a known allergy to bupivacaine were excluded. The target sample size was 40 patients; however, to enhance reliability, a total of 50 patients were enrolled.

Study Design Declaration

The initial ethics-approved protocol was designed as a randomized comparison between subcostal TAP block and port-site local infiltration. However, due to limited patient recruitment, randomization could not be executed. This report therefore represents an observational analysis of patients who received the surgeon-performed, laparoscopic-guided subcostal TAP block in accordance with the originally approved protocol. No additional procedures, interventions, or deviations from the ethics approval were undertaken.

Ethical Considerations

This study was approved by the Institutional Ethics Committee of Srinakharinwirot University (ethics code: SWUEC-004/2566F). Written informed consent was obtained from all participants prior to enrollment, and all patient data were collected and managed in accordance with institutional and international standards for data confidentiality and ethical research practice.

The study was retrospectively registered with the Thai Clinical Trials Registry (TCTR20250314002) following a change in study execution from the originally approved randomized protocol to a prospective observational design. Importantly, no protocol deviations occurred beyond the scope approved by the institutional ethics committee.

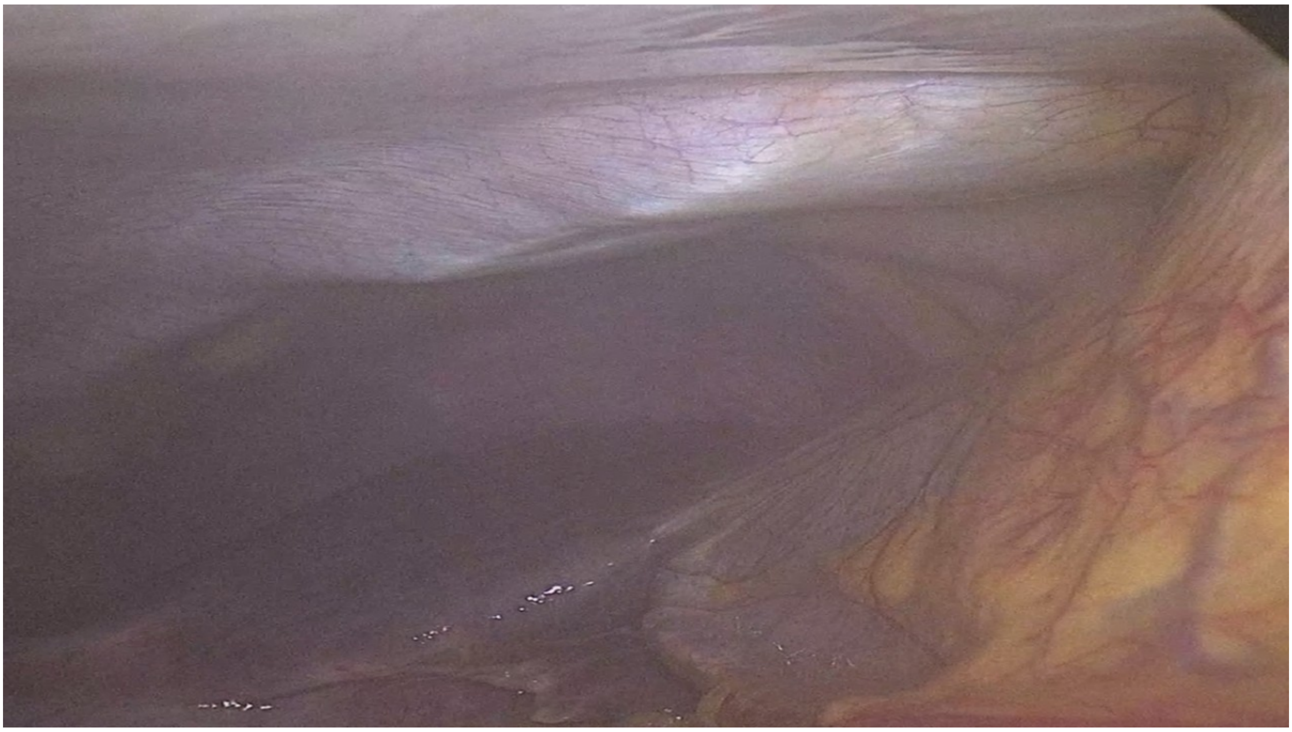
Laparoscopic-Guided Subcostal TAP Block Technique

LC was performed under general anesthesia. Intraoperative analgesia was standardized according to the institutional protocol and was not analyzed separately. Prior to skin incision, 5 mL of 0.25% bupivacaine was infiltrated at the umbilical site for local analgesia. A 12-mm umbilical port was inserted for the laparoscopic camera, and pneumoperitoneum was established at an intra-abdominal pressure of 8 to 12 mmHg.

After establishment of pneumoperitoneum, the right upper quadrant was inspected under laparoscopic visualization. The injection point was identified at the right subcostal region, approximately 2 to 3 cm inferior to the costal margin and lateral to the midline, corresponding to the junction between the posterior rectus sheath and the transversus abdominis muscle.

Under direct laparoscopic visualization, a long spinal needle was inserted percutaneously toward the TAP, with initial advancement at an angle of approximately 60 to 80 degrees relative to the abdominal wall. Upon reaching the target fascial plane, the needle angle was adjusted to approximately 30 degrees to facilitate controlled anesthetic delivery. After negative aspiration, 20 mL of 0.25% bupivacaine was injected incrementally. Correct local anesthetic deposition was inferred from laparoscopic visualization of the Doyle bulge, indicating separation of the posterior rectus sheath and transversus abdominis muscle as a surrogate marker of appropriate fascial plane injection (Figure 1).

Figure 1. Illustration of Doyle bulge observed during laparoscopic-guided subcostal transversus abdominis plane block.



In our contemporary LC practice, routine use of a 12-mm epigastric working port has been replaced by three 5-mm working ports positioned along the right subcostal margin, in addition to the 12-mm umbilical optical port. This port configuration reflects an evolving minimally invasive approach aimed at reducing port-site trauma while maintaining adequate operative ergonomics and visualization. After cholecystectomy, the gallbladder was retrieved in a sterile bag through the 12-mm umbilical port. A supplementary video demonstrating the laparoscopic-guided subcostal TAP block technique is provided in Multimedia Appendix 1.

Postoperative Pain Management and Monitoring

All patients received standardized postoperative analgesia. Oral acetaminophen (500 mg every 6 hours) and naproxen (250 mg twice daily) were prescribed as first-line agents; tramadol (50 mg twice daily) was substituted for patients with nonsteroidal anti-inflammatory drug intolerance. Persistent pain with a Visual Analog Scale (VAS) score >5 despite oral medication was managed with intravenous morphine (0.1 mg/kg every 4 hours as needed). Ondansetron (4 mg every 8 hours as required) was given for nausea and vomiting prophylaxis.

Pain intensity was evaluated using the VAS (0=no pain and 10=worst pain) at 2, 4, 6, 8, 12, and 24 hours postoperatively and during mobilization. Total morphine consumption and opioid-related adverse events (nausea, vomiting, and urinary retention) were recorded within the first 24 hours. Patients were observed for possible bupivacaine toxicity during this period.

Clinical Variables

Data collection included patient demographics, comorbidities, laboratory parameters, operative details, postoperative outcomes, and histopathological findings. Laboratory parameters analyzed in this study were preoperative values obtained within 24 hours before surgery. Operative variables comprised surgical duration, estimated blood loss, and intraoperative complications.

Postoperative variables included pain scores, total morphine consumption, and postoperative complications. All data were prospectively recorded in a predesigned spreadsheet (Microsoft Excel) for statistical analysis. The dataset was stored in password-protected files on secure institutional computers, with access restricted to study investigators only. Patient identifiers were removed and replaced with coded study numbers to ensure confidentiality.

Statistical Analysis

Statistical analysis was conducted using SPSS Statistics (version 27.0; IBM Corp). Descriptive statistics (mean, SD, frequencies, and percentages) were used to summarize demographic and clinical data. Comparisons between subgroups were conducted using the chi-square or Fisher exact test for categorical variables and the independent 2-tailed *t* test (or 1-way ANOVA where applicable) for continuous variables.

Exploratory multivariable logistic regression was applied to assess potential associations between perioperative factors and postoperative morphine requirement, acknowledging the limited sample size. Variables included in the multivariable model were selected based on clinical relevance and prior literature, rather than solely on statistical significance in univariate analyses. Given the limited number of outcome events, the analysis was conducted with a restricted number of covariates to maintain an acceptable events per variable ratio and minimize the risk of overfitting. A 2-tailed *P* value <.05 was considered statistically significant.

Results

Overview

During the study period, 50 patients underwent LC. In total, 8 (16%) patients were excluded due to conversion to open surgery, leaving 42 (84%) patients for analysis. Of these, 21 (50%) patients required postoperative morphine within the first 24 hours. Among patients who received opioids (n=21, 50%), the mean cumulative morphine dose was 3.86 (SD 1.39) mg.

Patient Characteristics

Baseline demographic and clinical characteristics are summarized in Table 1. There were no significant differences between the morphine-required and morphine-free groups in age (*P*=.55), BMI (*P*=.55), or sex distribution (*P*=.22). Comorbid conditions were more frequent in the morphine-required group; however, this difference did not reach statistical significance (*P*=.10). The mean American Society of Anesthesiologists (ASA) classification was significantly higher among patients who required morphine (mean 2.14, SD 0.57 vs mean 1.67, SD 0.73; *P*=.024), reflecting a higher baseline perioperative risk profile.

Table . Baseline demographic and clinical characteristics of the study population.

Variable	Morphine required (n=21)	Morphine free (n=21)	*P* value
Age (years), mean (SD)	57.1 (14.1)	54.3 (15.8)	.55
Sex (female), n (%)	13 (61.9)	8 (38.1)	.22
BMI (kg/m²), mean (SD)	25.8 (5.3)	24.8 (4.9)	.55
ASA[a] classification, mean (SD)	2.14 (0.57)	1.67 (0.73)	.02
Any comorbidity, n (%)	17 (81.0)	10 (47.6)	.10
Diabetes mellitus, n (%)	10 (47.6)	6 (28.6)	.21
Hypertension, n (%)	13 (61.9)	10 (47.6)	.36
Cardiovascular disease, n (%)	6 (28.6)	4 (19.0)	.48
Indication for surgery, n (%)[b]			
Symptomatic gallstone	17 (81.0)	15 (71.4)	.48
Acute cholecystitis	2 (9.5)	4 (19.0)	.39
Interval LC[c] after conservative treatment of acute cholecystitis	4 (19.0)	4 (19.0)	>.99
Previous ERCP[d], n (%)	6 (28.6)	5 (23.8)	.73

[a]ASA: American Society of Anesthesiologists.

[b]Categories are not mutually exclusive.

[c]LC: laparoscopic cholecystectomy.

[d]ERCP: endoscopic retrograde cholangiopancreatography.

Preoperative Laboratory Parameters

Preoperative laboratory findings are presented in Table 2. All laboratory values represent measurements obtained within 24 hours prior to surgery. Patients in the morphine-required group had significantly lower baseline hemoglobin (*P*=.01) and hematocrit levels (*P*=.002). Additionally, no significant between-group differences were observed in white blood cell count, neutrophil to lymphocyte ratio, liver function tests, or serum albumin levels.

Table . Preoperative laboratory parameters.

Variable	Morphine required (n=21), mean (SD)	Morphine free (n=21), mean (SD)	*P* value
Hemoglobin (g/dL)	11.95 (1.78)	13.34 (1.60)	.01
Hematocrit (%)	35.9 (4.8)	40.2 (4.0)	.002
White blood cell count (×10³/μL)	11.6 (19.1)	10.7 (11.1)	.85
Neutrophil to lymphocyte ratio	6.92 (13.8)	4.54 (6.54)	.48
Aspartate aminotransferase (U/L)	41.2 (54.4)	45.2 (55.5)	.81
Alanine aminotransferase (U/L)	38.8 (63.7)	46.0 (69.6)	.73
Alkaline phosphatase (U/L)	83.8 (41.2)	95.1 (84.9)	.58
Total bilirubin (mg/dL)	0.70 (0.51)	0.99 (1.30)	.35
Serum albumin (g/dL)	4.33 (0.39)	4.21 (0.34)	.56

Operative Details and Postoperative Outcomes

Operative and postoperative outcomes are summarized in Table 3. Mean operative time (mean 63.5, SD 15.6 minutes vs mean 58.4, SD 18.0 minutes; *P*=.32), estimated blood loss (mean 15.7, SD 11.3 mL vs mean 13.3, SD 12.5 mL; *P*=.52), and length of hospital stay (mean 2.57, SD 0.98 days vs mean 2.33, SD 0.86 days; *P*=.40) did not differ significantly between groups. Furthermore, no significant differences were observed between groups regarding trocar placement, operative technique, operative time, or intraoperative complications, and background analgesic regimens were comparable. Although patients with gallstone-related complications were included as indications for surgery, final histopathological examination of gallbladder specimens was reported as acute or chronic cholecystitis.

Table . Operative details, analgesic regimen, and postoperative outcomes.

Variable	Morphine required (n=21)	Morphine free (n=21)	*P* value
Analgesic regimen			
Acetaminophen+naproxen, n (%)	7 (33.3)	12 (57.1)	.12
Acetaminophen+tramadol, n (%)	14 (66.7)	9 (42.9)	.22
Operative details			
Complete cholecystectomy, n (%)	20 (95.2)	19 (90.5)	.56
Partial cholecystectomy, n (%)	1 (4.7)	2 (9.5)	.56
Operative time (minutes), mean (SD)	63.5 (15.6)	58.4 (18.0)	.32
Estimated blood loss (mL), mean (SD)	15.7 (11.3)	13.3 (12.5)	.52
Postoperative outcomes			
Postoperative complication, n (%)	1 (4.7)	1 (4.7)	.44
Nausea and vomiting, n (%)	1 (4.7)	1 (4.7)	.44
Hospital stay (days), mean (SD)	2.57 (0.98)	2.33 (0.86)	.40
Pathology: chronic cholecystitis, n (%)	20 (95.3)	17 (81.0)	.16
Pathology: acute cholecystitis, n (%)	1 (4.7)	4 (19.0)	.30

Factors Associated With Postoperative Morphine Requirement

Factors associated with postoperative morphine requirement in exploratory multivariable analysis are summarized in Table 4. Higher ASA class was associated with increased odds of morphine use within 24 hours after surgery (*P*=.01; odds ratio 6.51, 95% CI 1.37 - 30.96). Lower hemoglobin level demonstrated a trend toward association with morphine requirement but did not reach statistical significance (*P*=.07; odds ratio 0.58, 95% CI 0.32 - 1.06). Other variables, including age, sex, gallstone-related complications, and history of endoscopic retrograde cholangiopancreatography, were not significantly associated with opioid use.

Table . Multivariable logistic regression analysis of factors associated with postoperative morphine requirement.

Variable	Odds ratio (95% CI)	*P* value
Female	1.81 (0.26-12.80)	.55
ASA[a] classification	6.51 (1.37-30.96)	.01
Hemoglobin (per g/dL)	0.58 (0.32-1.06)	.07
Age (years)	0.98 (0.92-1.04)	.46
Gallstone-related complication	0.91 (0.16-5.15)	.91
Previous ERCP[b]	0.75 (0.12-4.56)	.75

[a]ASA: American Society of Anesthesiologists.

[b]ERCP: endoscopic retrograde cholangiopancreatography.

Given the limited number of outcome events, regression analyses were conducted within an exploratory framework constrained by events per variable considerations.

Postoperative Pain Scores

Postoperative pain scores are summarized in Table 5. Patients who required morphine reported higher VAS scores during the early postoperative period, particularly at 2 hours (mean 3.29, SD 1.45 vs mean 1.93, SD 0.96; *P*=.009), 4 hours (mean 3.95, SD 1.39 vs mean 2.07, SD 1.00; *P*<.001), and 12 hours (mean 3.57, SD 1.44 vs mean 2.43, SD 1.06; *P*=.02). At 6 hours, 24 hours, and during mobilization, pain scores remained numerically higher in the morphine-required group but did not reach statistical significance.

Table . Comparison of postoperative pain scores between morphine-required and morphine-free groups.

Time point	Morphine required, mean (SD)	Morphine free, mean (SD)	*P* value
2 hours	3.29 (1.45)	1.93 (0.96)	.009
4 hours	3.95 (1.39)	2.07 (1.00)	<.001
6 hours	3.33 (1.43)	2.73 (1.42)	.26
12 hours	3.57 (1.44)	2.43 (1.06)	.02
24 hours	2.57 (1.21)	1.97 (0.81)	.07
During mobilization	4.29 (1.33)	3.97 (1.25)	.40

These findings descriptively reflect differences in pain experience between groups and are presented to contextualize postoperative opioid requirement rather than to infer comparative analgesic effectiveness.

Discussion

Principal Findings

This prospective observational study suggests a clinically relevant opioid-sparing association of laparoscopic-guided subcostal TAP block in patients undergoing LC. Approximately half of the patients (n=21, 50%) did not require postoperative opioids, while those who did (n=21, 50%) received only a modest cumulative dose (mean 3.86, SD 1.39 mg).

The observed analgesic pattern was most evident during the early postoperative period, particularly between 2 and 12 hours, which is consistent with the expected pharmacodynamic profile of 0.25% bupivacaine. Within the context of this observational cohort, these findings support the feasibility of surgeon-performed TAP block as a practical adjunct to multimodal analgesia strategies, with the potential to limit postoperative opioid exposure while maintaining adequate pain control [11-14].

The results of this study are directionally consistent with previous reports suggesting that thoracoabdominal and subcostal TAP blocks are associated with improved early postoperative pain control following LC [15-17]. Importantly, this study extends the existing literature by including patients with complicated gallstone disease, a population that has been relatively underrepresented in prior research.

From a mechanistic perspective, the observed analgesic association may be attributable to localized somatic blockade of the upper abdominal wall corresponding to trocar insertion sites. Such coverage is thought to attenuate incisional and parietal peritoneal pain, which may explain the more pronounced pain relief observed during the first 12 postoperative hours [10,18]. Clinically, these observations underscore the importance of integrating regional analgesic techniques with scheduled nonopioid coanalgesics and appropriately timed rescue analgesia to maintain adequate pain control within ERAS-oriented perioperative pathways [5,6,13,14].

Exploratory predictor analysis suggested that patient-related factors were more strongly associated with postoperative opioid requirement than intraoperative variables. Higher ASA classification was independently associated with postoperative morphine use, indicating that greater comorbidity burden may be linked to increased analgesic needs despite regional blockade.

These findings support a risk-adapted approach to perioperative pain management for patients at higher risk.

Although the TAP block primarily targets somatic abdominal wall pain, unmeasured factors such as visceral pain burden, neuropathic pain components, and subtle variations in block accuracy may have influenced postoperative pain perception and opioid requirement. These factors were not objectively assessed in this study and may contribute to residual variability beyond patient-level characteristics such as ASA classification.

Beyond its analgesic association, surgeon-performed subcostal TAP block offers several practical advantages, particularly in settings with limited anesthesiology support or restricted access to ultrasound equipment. Incorporation of this technique into the laparoscopic workflow allows a consistent and equipment-independent approach to regional analgesia, aligning with broader initiatives in opioid stewardship and sustainable perioperative care [9,15-17,19].

Several limitations should be acknowledged. The single-center, nonrandomized design limits causal inference, and the absence of a control group precludes direct comparison with standard port-site local anesthetic infiltration. The modest sample size restricts statistical precision; therefore, multivariable analyses were conducted within an exploratory framework with selective variable inclusion to reduce the risk of overfitting.

In addition, no formal postoperative sensory testing was performed to objectively verify block success, primarily due to practical considerations within the perioperative workflow and the study's focus on clinically relevant outcomes. Instead, correct local anesthetic deposition was inferred from laparoscopic visualization of the Doyle bulge, indicating separation of the posterior rectus sheath and transversus abdominis muscle as a surrogate marker of appropriate fascial plane injection [20,21]. However, this visual confirmation cannot fully substitute for objective sensory testing and does not confirm the extent or consistency of dermatomal blockade.

Substitution of tramadol for nonsteroidal anti-inflammatory drugs in a small number of patients may have introduced minor confounding, although perioperative analgesic protocols were otherwise standardized. In addition, psychosocial factors known to influence postoperative pain, including anxiety, pain catastrophizing, and prior opioid exposure, were not assessed [22,23]. These limitations should be considered when interpreting the findings.

Despite these constraints, this study has several notable strengths. The prospective data collection, use of standardized perioperative analgesic pathways, and inclusion of patients with both uncomplicated and complicated gallstone disease enhance the clinical relevance of the study. Overall, the findings indicate that surgeon-performed, laparoscopic-guided subcostal TAP block is a technically straightforward and reproducible adjunct within multimodal analgesia strategies. The observed early opioid-sparing association supports the feasibility of considering this technique in ERAS-oriented perioperative pathways and broader opioid reduction efforts.

The relatively longer hospital stay observed in this cohort reflects local institutional practice, in which LC is not routinely performed as a day-case procedure. Inclusion of patients with complicated gallstone disease required preoperative admission and postoperative observation for safety. In addition, routine preoperative admission at least 1 day prior to surgery, according to institutional protocol, contributed to the overall length of hospital stay.

Future research should include multicenter randomized controlled trials comparing laparoscopic-guided TAP block with standard port-site local anesthetic infiltration. On the basis of the observed effect estimates, a sample size of approximately 80 to 100 patients per arm may provide 80% power to detect a 1-point difference in mean VAS pain score at an α level of .05. Incorporation of cost-effectiveness analyses, patient-reported outcome measures, and longer-term follow-up would further clarify the clinical value, scalability, and role of this technique in minimally invasive surgery.

Conclusions

This study suggests that laparoscopic-guided subcostal TAP block may be associated with lower early postoperative pain scores and reduced opioid requirements in patients undergoing LC, including those with gallstone-related complications. Although limited by sample size, the findings support the feasibility of considering surgeon-performed subcostal TAP block as part of multimodal analgesia strategies within ERAS-oriented perioperative pathways.

Acknowledgments

The authors thank the surgical staff at the Department of Surgery, Faculty of Medicine, Srinakharinwirot University, for their assistance with data acquisition. During the preparation of this work, the authors used ChatGPT to improve grammar and language clarity. After using this tool, the authors carefully reviewed, revised, and approved the content and take full responsibility for the published article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

TT contributed to the conception and design of the study, data acquisition, data analysis and interpretation, and manuscript drafting. SM contributed to study design, data collection and analysis, and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Laparoscopic-guided subcostal transversus abdominis plane (TAP block).

[MP4 File, 98727 KB - periop_v9i1e87622_app1.mp4]

References

1. Tullavardhana T. Critical view of safety: a safe method to prevent bile duct injury from laparoscopic cholecystectomy. J Med Health Sci 2015;22:49-57 [FREE Full text]
2. Keir A, Rhodes L, Kayal A, Khan OA. Does a transversus abdominis plane (TAP) local anaesthetic block improve pain control in patients undergoing laparoscopic cholecystectomy? A best evidence topic. Int J Surg 2013;11(9):792-794. [doi: 10.1016/j.ijsu.2013.05.039] [Medline: 23770342]
3. Law LS, Lo EA, Gan TJ, et al. Preoperative antiemetic and analgesic management. In: Gan TJ, Thacker JK, Miller TE, editors. Enhanced Recovery for Major Abdominopelvic Surgery: Professional Communications Inc; 2016:105-120.
4. Ljungqvist O, Scott M, Fearon KC. Enhanced recovery after surgery: a review. JAMA Surg 2017 Mar 1;152(3):292-298. [doi: 10.1001/jamasurg.2016.4952] [Medline: 28097305]
5. Kaye AD, Urman RD, Rappaport Y, et al. Multimodal analgesia as an essential part of enhanced recovery protocols in the ambulatory settings. J Anaesthesiol Clin Pharmacol 2019 Apr;35(Suppl 1):S40-S45. [doi: 10.4103/joacp.JOACP_51_18] [Medline: 31142958]
6. Hebbard P, Fujiwara Y, Shibata Y, Royse C. Ultrasound-guided transversus abdominis plane (TAP) block. Anaesth Intensive Care 2007 Aug;35(4):616-617. [doi: 10.1177/0310057X0703500417] [Medline: 18020088]
7. Wang W, Wang L, Gao Y. A meta-analysis of randomized controlled trials concerning the efficacy of transversus abdominis plane block for pain control after laparoscopic cholecystectomy. Front Surg 2021;8:700318. [doi: 10.3389/fsurg.2021.700318] [Medline: 34422893]
8. Elamin G, Waters PS, Hamid H, et al. Efficacy of a laparoscopically delivered transversus abdominis plane block technique during elective laparoscopic cholecystectomy: a prospective, double-blind randomized trial. J Am Coll Surg 2015 Aug;221(2):335-344. [doi: 10.1016/j.jamcollsurg.2015.03.030] [Medline: 25899736]
9. Venkatraman R, Saravanan R, Dhas M, Pushparani A. Comparison of laparoscopy-guided with ultrasound-guided subcostal transversus abdominis plane block in laparoscopic cholecystectomy - a prospective, randomised study. Indian J Anaesth 2020 Dec;64(12):1012-1017. [doi: 10.4103/ija.IJA_528_20] [Medline: 33542563]
10. Alsharari AF, Abuadas FH, Alnassrallah YS, Salihu D. Transversus abdominis plane block as a strategy for effective pain management in patients with pain during laparoscopic cholecystectomy: a systematic review. J Clin Med 2022 Nov 22;11(23):6896. [doi: 10.3390/jcm11236896] [Medline: 36498471]
11. Simpson JC, Bao X, Agarwala A. Pain Management in Enhanced Recovery after Surgery (ERAS) protocols. Clin Colon Rectal Surg 2019 Mar;32(2):121-128. [doi: 10.1055/s-0038-1676477] [Medline: 30833861]
12. Cao L, Yang T, Hou Y, Yong S, Zhou N. Efficacy and safety of different preemptive analgesia measures in pain management after laparoscopic cholecystectomy: a systematic review and network meta-analysis of randomized controlled trials. Pain Ther 2024 Dec;13(6):1471-1497. [doi: 10.1007/s40122-024-00647-w] [Medline: 39227523]
13. Barazanchi AWH, MacFater WS, Rahiri JL, et al. Evidence-based management of pain after laparoscopic cholecystectomy: a PROSPECT review update. Br J Anaesth 2018 Oct;121(4):787-803. [doi: 10.1016/j.bja.2018.06.023] [Medline: 30236241]
14. Cho HY, Hwang IE, Lee M, Kwon W, Kim WH, Lee HJ. Comparison of modified thoracoabdominal nerve block through perichondral approach and subcostal transversus abdominis plane block for pain management in laparoscopic cholecystectomy: a randomized-controlled trial. Korean J Pain 2023 Jul 1;36(3):382-391. [doi: 10.3344/kjp.23114] [Medline: 37336756]
15. Tolchard S, Davies R, Martindale S. Efficacy of the subcostal transversus abdominis plane block in laparoscopic cholecystectomy: Comparison with conventional port-site infiltration. J Anaesthesiol Clin Pharmacol 2012 Jul;28(3):339-343. [doi: 10.4103/0970-9185.98331] [Medline: 22869941]
16. Ozciftci S, Sahiner Y, Sahiner IT, Akkaya T. Is Right Unilateral Transversus Abdominis Plane (TAP) block successful in postoperative analgesia in laparoscopic cholecystectomy? Int J Clin Pract 2022;2022:2668215. [doi: 10.1155/2022/2668215] [Medline: 35685608]
17. Emile SH, Elfeki H, Elbahrawy K, Sakr A, Shalaby M. Ultrasound-guided versus laparoscopic-guided subcostal transversus abdominis plane (TAP) block versus No TAP block in laparoscopic cholecystectomy; a randomized double-blind controlled trial. Int J Surg 2022 May;101:106639. [doi: 10.1016/j.ijsu.2022.106639] [Medline: 35487422]
18. Hinrichs-Rocker A, Schulz K, Järvinen I, Lefering R, Simanski C, Neugebauer EAM. Psychosocial predictors and correlates for chronic post-surgical pain (CPSP) - a systematic review. Eur J Pain 2009 Aug;13(7):719-730. [doi: 10.1016/j.ejpain.2008.07.015] [Medline: 18952472]

19. Boonyapalanant C, Woranisarakul V, Jitpraphai S, et al. The efficacy of inside-out transversus abdominis plane block vs local infiltration before wound closure in pain management after kidney transplantation: a double-blind, randomized trial. Siriraj Med J 2022;74(4):233-238. [doi: 10.33192/Smj.2022.29]
20. Vindal A, Sarda H, Lal P. Laparoscopically guided transversus abdominis plane block offers better pain relief after laparoscopic cholecystectomy: results of a triple blind randomized controlled trial. Surg Endosc 2021 Apr;35(4):1713-1721. [doi: 10.1007/s00464-020-07558-9] [Medline: 32297054]
21. Suragul W, Tantawanit A, Rungsakulkij N, et al. Effect of local anaesthetic infiltration on postoperative pain after laparoscopic cholecystectomy: randomized clinical trial. BJS Open 2022 May 2;6(3):zrac066. [doi: 10.1093/bjsopen/zrac066] [Medline: 35639946]
22. Ip HYV, Abrishami A, Peng PWH, Wong J, Chung F. Predictors of postoperative pain and analgesic consumption: a qualitative systematic review. Anesthesiology 2009 Sep;111(3):657-677. [doi: 10.1097/ALN.0b013e3181aae87a] [Medline: 19672167]
23. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res 2010 Mar;468(3):798-806. [doi: 10.1007/s11999-009-0963-y] [Medline: 19585177]

Abbreviations

ERAS: enhanced recovery after surgery
LC: laparoscopic cholecystectomy
TAP: transversus abdominis plane
VAS: Visual Analog Scale

Edited by N Rohatgi, S Pan; submitted 11.Nov.2025; peer-reviewed by A Jahangeer, ADK Hill, H Abdallah; revised version received 22.Feb.2026; accepted 25.Feb.2026; published 25.Mar.2026.

Please cite as:
Mahasupachai S, Tullavardhana T
Clinical Feasibility and Outcomes of Surgeon-Performed Laparoscopic-Guided Subcostal Transversus Abdominis Plane Block in Laparoscopic Cholecystectomy: Prospective Observational Study
JMIR Perioper Med 2026;9:e87622
URL: https://periop.jmir.org/2026/1/e87622
doi:10.2196/87622

Assessing the Effects of eHealth Literacy and the Area Deprivation Index on Barriers to Electronic Patient Portal Use for Orthopedic Surgery: Cross-Sectional Observational Study

Audrey Lynn Litvak[1*], BA; Nicholas Lin[2*], MD; Kelly Hynes[3*], MD; Jason Strelzow[3*], MD; Jeffrey G Stepan[2*], MD, MSc

[1]Pritzker School of Medicine, University of Chicago, 924 E 57th Street #104, Chicago, IL, United States

[2]Department of Orthopaedic Surgery and Rehabilitation Services, University of Chicago, Chicago, IL, United States

[3]Department of Orthopaedic Surgery, Washington University in St Louis, St Louis, MO, United States

[*]all authors contributed equally

Corresponding Author:
Audrey Lynn Litvak, BA
Pritzker School of Medicine, University of Chicago, 924 E 57th Street #104, Chicago, IL, United States

Abstract

Background: As electronic patient portals (EPPs) continue to gain popularity and systems transition to online tools for scheduling, communication, and telehealth, patients without access or skills to use these tools may be overlooked.

Objective: This study analyzed patient and neighborhood-level factors, including eHealth literacy level and the Area Deprivation Index (ADI), that may limit EPP access for orthopedic surgery.

Methods: A cross-sectional, survey-based study was performed at a single urban tertiary academic medical center in the United States across foot and ankle, hand and upper extremity, and orthopedic trauma subspecialty clinics from June 21, 2022, to August 12, 2022. Survey responses (N=287) provided information on sociodemographic characteristics; barriers to EPP use and frequency of EPP use; the eHealth Literacy Scale; and the ADI, which is an address-generated national census measure of neighborhood-level disadvantage. Barriers to EPP use were inductively coded into barrier types, classified as physical access, technology discomfort, or preference. The primary outcome measure was patient-reported barriers to EPP use, which was treated as a binary outcome (1=barrier; 0=no barrier). Bivariate analyses and multivariable binary logistic regressions were performed.

Results: The percentage of patients who self-reported barriers to EPP access was 43.2% (124/287), which related to physical access (13/124, 10.4%), technology discomfort (55/124, 44.3%), and preference (78/124, 63.0%). In the adjusted regressions, only low eHealth literacy and older age predicted barriers to EPP use (low eHealth literacy, adjusted odds ratio [AOR] 1.32, 95% CI 1.13-1.54; *P*<.001; older age, AOR 1.007, 95% CI 1.003-1.009; *P*<.001), including barriers of technology discomfort (low eHealth literacy, AOR 1.25, 95% CI 1.11-1.40; *P*<.001; older age, AOR 1.004, 95% CI 1.002-1.007; *P*<.001) and preference (low eHealth literacy, AOR 1.33, 95% CI 1.17-1.51; *P*<.001; older age, AOR 1.004, 95% CI 1.00-1.01; *P*<.01). Patients with physical access–related barriers as opposed to technology discomfort or preference barriers had the lowest median eHealth literacy scores (17.0, IQR 12.0-14.0 vs 27.0, IQR 16.0-32.0 vs 27.0, IQR 20.0-32.0, respectively) and roughly a quartile higher median ADI (73.0, IQR 41.0-92.0 vs 53.5, IQR 31.2-76.0 vs 58.0, IQR 38.8-83.8, respectively).

Conclusions: Low eHealth literacy was the most significant determinant of overall barriers to EPP use for orthopedic surgery, followed by older age. Neighborhood-level disadvantage as measured through the ADI had no mediating effect on patient-reported barriers to EPP use when adjusting for eHealth literacy level. While patients with physical access barriers had higher ADIs, overall, few patients reported physical access barriers compared to barriers related to technology discomfort or preference. Patient preference for EPP versus non-EPP communications should be documented. Point-of-care screening using the eHealth Literacy Scale may also identify patients who require follow-up outside of the EPP during critical perioperative periods.

(JMIR Perioper Med 2026;9:e72035) doi:10.2196/72035

KEYWORDS

electronic health records; eHealth literacy; online systems; health equity; social determinants of health; SDOH; orthopedic surgery

Introduction

Improving digital health information transparency and transmission via electronic patient portals (EPPs) has been a central focus of health IT policy in the United States during the last decade [1]. EPPs facilitate patients' communication with their treating care teams and direct access to their electronic personal health record and tools to request prescription refills, participate in e-visits, and complete patient-reported outcome questionnaires, among other functions. These interactions empower patients to take an active role in their health care [1,2]. However, few studies have examined patient portal use in the surgical setting or patient factors that may limit EPP use in this context [3-8].

Benefits of EPP enrollment among orthopedic patients include improved patient outcomes [4,7], medication adherence [3], higher patient satisfaction and psychosocial health [3,7], and fewer missed appointments [3]. Patient engagement via the EPP may additionally facilitate more effective screening for commonly avoidable complications that delay patients' return to function, such as soaking of splints or patient-prolonged immobilization due to unanticipated postoperative pain. Moreover, as routine messaging and completion of patient-reported outcomes via the EPP becomes standard, it is likely that patient engagement via the EPP beyond enrollment may become another critical quality metric tied to physician reimbursement.

Despite advantages of and health care provider interest in adoption of EPP tools, prior studies in orthopedic surgery have shown that patient factors, including older age and lower educational level, may limit EPP enrollment [3,4,6,8], which is analogous to observations in the internal medicine setting [9-12]. The 2020 Health Information National Trends Survey (HINTS) found that the most commonly cited reason for patient nonuse within a large US sample was desire to speak directly with a health care provider (ie, physician, nurse practitioner)—a sentiment shared by 69% of patients [11]. Furthermore, roughly 30% of patients expressed discomfort with the technology [11]. Traditional health literacy refers to the capacity to find, understand, and use health information to inform health-related decisions and actions, whereas eHealth literacy specifically refers to the capacity and skills to seek, assess, and make use of health information via electronic media. In the hospital medicine setting, low eHealth literacy in particular is associated with less awareness, use, and perceived usefulness of EPPs [13].

Lack of examination of granular patient-level use data beyond EPP activation status, explicit barriers to EPP use, and associated patient factors such as health literacy are described as significant limitations and directions for future work in orthopedic surgery [3,4,8]. Additionally, no study across any prior setting has assessed the effect of structural or neighborhood-level determinants on barriers to EPP use, nor have they assessed barriers among patients who are actively enrolled in EPPs. Individual-level determinants may refer to patient demographics or skill sets such as health literacy, whereas neighborhood-level determinants refer to unmeasured social factors conferred by the geographic environment in which a patient lives, often described via census variables related to percentage of unemployment, percentage of individuals with a high school education, and food and housing quality, among others.

For digital health uptake in particular, distinguishing among types and levels of determinants is critical to realizing equity-informed intervention and policy [14,15]. For example, digital literacy is an individual-level factor for which a policy-level solution such as improving broadband connectivity or personal device accessibility may be ineffective in the absence of community-responsive interventions to provide individuals with digital skill training [15]. In particular, it is important to analyze whether neighborhood disadvantage may amplify the impact of low eHealth literacy on barriers to patient portal use, which single-level analyses of eHealth literacy cannot capture.

This study aimed to contribute to the existing body of literature on barriers to EPP use in orthopedic surgery by analyzing how both individual-level and neighborhood-level social determinants, including eHealth literacy and the Area Deprivation Index (ADI), may relate to patient-reported barriers to EPP access and use across foot and ankle, hand and upper extremity, and orthopedic trauma surgery.

Methods

Study Design and Setting

This was a cross-sectional survey–based study conducted via an anonymized paper survey administered at a single urban tertiary academic medical center in the United States between June 21, 2022, and August 12, 2022. The survey was administered in the clinic following each patient visit and consisted of sociodemographic questions, the eHealth Literacy Scale (eHEALS), and 2 questions regarding EPP access and use detailed below. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) design and reporting guidelines for cross-sectional studies.

Participants

All English-speaking patients aged >18 years presenting for orthopedic surgery evaluation at foot and ankle, trauma, and hand and upper extremity clinics were included and approached consecutively during the aforementioned period. This study was limited to foot and ankle, trauma, and hand and upper extremity surgery clinics where faculty involvement at our institution was feasible. Approached patients were excluded if they had not received a tablet to complete their in-office patient-reported outcome questionnaire per routine standard of care due to external technological or capacity constraints (ie, tablet out of battery or too few working tablets in the clinic on a particular day) unless that patient had already completed the questionnaire via their patient portal. While this practical limitation resulted in potential selection bias, it randomly affected only a small portion of patients (<10) and was remedied to avoid recurrence of the problem for continued recruitment efforts. Patients who could not read or write were included and read aloud the study survey.

Sample Size Calculation

The HINTS 2020 reported that approximately 40% of adults in the United States accessed their patient portal in the previous year, whereas 59% were nonusers [11]. Using a baseline nonuse or potential barrier rate of roughly 40% to 60% and setting an α value of .05, we estimated that a sample of 270 to 290 patients would provide at least 80% power to detect large effect sizes (odds ratio 2.0 - 3.0) in a binary regression model constrained by 10 events per covariate included.

Ethical Considerations

Biological Sciences Division/University of Chicago Institutional Review Board approval (IRB22-0230) was obtained with waivers for written consent and HIPAA (Health Insurance Portability and Accountability Act) authorization. Anonymized survey data were transcribed into a REDCap (Research Electronic Data Capture; Vanderbilt University) database for secure storage [16]. No compensation was provided for participation.

Primary and Secondary Outcomes

The primary outcome measure was patient-reported barriers to EPP access. Barriers to access were derived from the HINTS and secondarily classified as barriers of physical access, discomfort with technology, or patient preferences for nonelectronic provider communication (Multimedia Appendix 1) [11]. These categories were inductively coded by the research team after data collection. This classification is not validated, which we discuss as a limitation. Patients were instructed to mark preference for nonelectronic provider communication as a selection only if they perceived their preferences as barriers to using their portals. Importantly, "I do not have a patient portal account" was listed as an option on the original survey but was analyzed separately. The secondary outcome measure was patient-reported level of EPP use classified into 2 categories: routine use and nonroutine use. Per Maroney et al [17], level of use was characterized as routine if at least monthly use was indicated and as nonroutine if use a few times a year or less frequently was indicated, including those who did not have an EPP (Multimedia Appendix 1). Importantly, the level of EPP use included use for any clinic, not limited only to their orthopedic surgery care.

Variables and Demographics

eHealth literacy was measured via patient responses to the eHEALS tool to determine its association with barriers to EPP use (Multimedia Appendix 2) [18]. This tool has been validated in the orthopedic outpatient setting, among others [18-20]. As in prior literature, a cumulative score of 25 or less indicated low eHealth literacy, and a score of 26 or greater indicated high eHealth literacy [13,21,22]. Neighborhood-level disadvantage was assessed using the ADI, which is calculated via publicly available census data in the domains of income, educational level, employment, and housing quality to assign numeric scores of societal disadvantage to particular geographical regions [23]. Higher scores indicate higher levels of societal disadvantage. Self-reported demographic data were additionally collected.

Statistical Analysis

Survey data were analyzed using the Python statistical program (version 3.10.6; Python Software Foundation) [24-26]. Missing values were excluded pointwise across the relevant analyses given the low frequency. Numerical data presented as medians were reported with the IQR. The level of significance was set at P=.05.

Bivariate analyses were performed to examine the association between both patient-reported barriers and level of EPP use and demographic variables, the ADI, and eHealth literacy level. Three multivariable logistic regressions were conducted wherein barriers were treated as a binary outcome (1=barrier; 0=no barrier). Categorical variables were converted to binary dummy variables for regression. The main regression considered all barriers, whereas the 2 subsequent regressions examined only barriers of technology discomfort or preference-related barriers. Regression was not performed for physical access–related barriers due to outcome size of 13, which is discussed as a limitation. Variable selection for each model was determined via outcome size (at least 10 outcomes per covariate to avoid overfitting) and one-at-a-time sensitivity analyses (via examination of the McFadden pseudo-R^2) in a backward regression approach. To avoid collinearity, only covariates with a variance inflation factor of <5 were included together. Model fit was assessed using McFadden pseudo-R^2 values, with 0.2 considered excellent if not overfit (corroborated via df).

Demographic categories and regression reference levels were selected based on breakdowns and historical controls used in prior related literature [9,27,28]. Income was treated categorically, and per the analogous literature, we selected 3 levels representative of low, medium, and high income based on the median household income cutoffs for the zip code tied to the authors' institution.

Results

Participant Characteristics

A convenience sample of 339 eligible patients was approached, of whom 52 (15.3%) declined participation, leaving 287 (84.7%) for analysis. The median age of the study participants was 48.5 (IQR 35.0-64.2) years; 58.2% (167/287) of the study participants self-identified as non-Hispanic Black individuals, and 26.1% (75/287) identified as non-Hispanic White individuals. The median cumulative eHEALS score was 32 (IQR 27-35), with 21.3% (61/287) of the study participants having low eHealth literacy (eHEALS score of 25 or less). The median ADI was 53.0 (IQR 32.0-74.8).

Of the study participants, 63.1% (181/287) were routine users, and only 9.8% (28/287) did not have EPPs. One or more barriers to accessing their EPPs were reported by 43.2% (124/287) of all patients. Moreover, among the 90.2% (259/287) of patients who were enrolled in the EPP, 42.5% (110/259) still reported barriers to access or use of their portal. The remaining patient characteristics were compared by self-reported barriers to EPP access and self-reported EPP use (Table 1). Patients reporting one or more barriers had higher median age than patients who did not report barriers (57.0, IQR 42.2-71.0 years vs 43.0, IQR

29.2-57.0 years; *P*<.001; Table 1). A higher percentage of non-Hispanic Black patients (*P*=.04), retirees (*P*=.001), and patients who fell into the income bracket of US $30,000 or less (*P*=.03) reported barriers to access (Table 1). Patients who did not routinely use the EPP had a lower educational level than routine users (*P*=.005; Table 1).

Table . Descriptive characteristics and exploratory comparison of self-reported barriers to electronic patient portal (EPP) access and self-reported EPP use by patient characteristics.[a]

Characteristic	Overall	No barriers	One or more barriers	*P* value	Routine use	Nonroutine use	*P* value
Subspecialty clinic, n/N (%)				.07			.77
Hand	187/287 (65.2)	111/187 (59.4)	76/187 (40.6)		116/187 (62.0)	71/187 (38.0)	
Foot and ankle	68/287 (23.7)	40/68 (58.8)	28/68 (41.2)		43/68 (63.2)	25/68 (36.8)	
Trauma	32/287 (11.1)	12/32 (37.5)	20/32 (62.5)		22/32 (68.8)	10/32 (31.3)	
Age (years), median (IQR)	48.5 (35.0-64.2)	43.0 (29.2-57.0)	57.0 (42.2-71.0)	<.001	47.0 (35.0-62.0)	52.0 (37.0-67.0)	.28
Race or ethnicity, n/N (%)				.04			.09
Hispanic or Latino	24/287 (8.4)	16/24 (66.7)	8/24 (33.3)		12/24 (50.0)	12/24 (50.0)	
Non-Hispanic Black	167/287 (58.2)	83/167 (49.7)	84/167 (50.3)		103/167 (61.7)	64/167 (38.3)	
Non-Hispanic White	75/287 (26.1)	51/75 (68.0)	24/75 (32.0)		48/75 (64.0)	27/75 (36.0)	
Other identity or preferred not to answer	21/287 (7.3)	13/21 (61.9)	8/21 (38.0)		18/21 (85.7)	3/21 (14.3)	
Highest educational level attained, n/N (%)				.07			.005
High school or lower	82/284 (28.9)	39/82 (47.6)	43/82 (52.4)		41/82 (50.0)	41/82 (50.0)	
Some college	73/284 (25.7)	40/73 (54.8)	33/73 (45.2)		45/73 (61.6)	28/73 (38.4)	
College or higher	129/284 (45.4)	82/129 (63.6)	47/129 (36.4)		93/129 (72.1)	36/129 (27.9)	
Annual income bracket (US $), n/N (%)				.03			.09
≤30,000	89/260 (34.2)	42/89 (47.2)	47/89 (52.8)		52/89 (58.4)	37/89 (41.6)	
30,001-50,000	48/260 (18.5)	26/48 (54.2)	22/48 (45.8)		30/48 (62.5)	18/48 (37.5)	
>50,000	123/260 (47.3)	80/123 (65.0)	43/123 (35.0)		89/123 (72.4)	34/123 (27.6)	
Current employment status, n/N (%)				.001			.19
Employed	160/285 (56.1)	105/160 (65.6)	55/160 (34.4)		109/160 (68.1)	51/160 (31.9)	
Unemployed or on disability	66/285 (23.2)	35/66 (53.0)	31/66 (47.0)		38/66 (57.6)	28/66 (42.4)	
Retired	59/285 (20.7)	22/59 (37.3)	37/59 (62.7)		34/59 (57.6)	25/59 (42.4)	

[a]Percentages may not add up to 100 due to rounding.

Exploratory Analysis of eHealth and ADI as Potential Barriers to EPP Access and Use

In the analysis of eHealth literacy, patients who reported barriers had a lower median eHEALS score than patients who did not report barriers (29.0, IQR 22.0-32.2 vs 32.0, IQR 30.0-38.0; *P*<.001; Table 2). Conversely, a higher percentage of patients with low eHealth literacy compared to high eHealth literacy reported barriers to using their EPPs (45/61, 74% vs 79/224, 35%, respectively; *P*<.001; Table 2). Patients who reported barriers also had higher median national ADI than patients who did not report barriers (55.5, IQR 37.5-78.2 vs 51.0, IQR 32.0-70.0; *P*=.06), and a higher percentage of patients from the most deprived ADI quartile also reported barriers compared to patients from the least deprived ADI quartiles (most deprived ADI quartile [76-100], 37/70, 53% vs second most deprived ADI quartile [51-75], 35/87, 40% vs second least deprived ADI quartile [26-50], 29/70, 41% vs least deprived ADI quartile [1-25], 19/55, 35%; *P*=.19); however, these results did not reach statistical significance (Table 2).

Table . Comparison of self-reported barriers to electronic patient portal access by eHealth literacy level and the Area Deprivation Index (ADI).[a]

	Overall	No barriers	Any barriers	*P* values
eHEALS[b] score, median (IQR)	32.0 (27.0-35.0)	32.0 (30.0-38.0)	29.0 (22.0-32.2)	<.001
eHealth literacy level (eHEALS score), n/N (%)				<.001
High eHealth literacy	224/285 (78.6)	145/224 (64.7)	79/224 (35.3)	
Low eHealth literacy	61/285 (21.4)	16/61 (26.2)	45/61 (73.8)	
National ADI, median (IQR)	53.0 (32.0-74.8)	51.0 (32.0-70.0)	55.5 (37.5-78.2)	.06
National ADI quartile, n/N (%)				.19
Least deprived ADI quartile (1-25)	55/282 (19.5)	36/55 (65.4)	19/55 (34.5)	
Second least deprived ADI quartile (26-50)	70/282 (24.8)	41/70 (58.6)	29/70 (41.4)	
Second most deprived ADI quartile (51-75)	87/282 (30.9)	52/87 (59.8)	35/87 (40.2)	
Most deprived ADI quartile (76-100)	70/282 (24.8)	33/70 (47.1)	37/70 (52.9)	

[a]Percentages may not add up to 100 due to rounding and missing values.
[b]eHEALS: eHealth Literacy Scale.

Most of the patient-reported barriers were related to technology discomfort (55/124, 44.4%) or preference (78/124, 62.9%) rather than physical access (13/124, 10.5%). However, the group reporting physical access barriers had the lowest levels of eHealth literacy (median eHEALS score 17.0 vs 27.0 vs 27.0, respectively) and highest ADI (median 73.0 vs 53.5 vs 58.0, respectively) compared to groups reporting barriers related to technology discomfort or preference (Table 3). Patients from the most deprived ADI quartile had higher percentages of barriers (Table 3). Bivariate analysis was not performed as barrier type was nonexclusive.

Table . Exploratory analysis of barrier type by eHealth literacy level and the Area Deprivation Index (ADI).[a]

	No barriers	Physical access barriers	Technology discomfort barriers	Preference barriers
eHEALS[b] score, median (IQR)	32.0 (30.0-38.0)	17.0 (12.0-24.0)	27.0 (16.0-32.0)	27.0 (20.0-32.0)
eHealth literacy level (eHEALS), n/N (%)				
High eHealth literacy	145/224 (64.7)	2/224 (0.9)	30/224 (13.4)	44/224 (19.6)
Low eHealth literacy	16/61 (26.2)	11/61 (18.0)	25/61 (41.0)	34/61 (55.7)
National ADI, median (IQR)	51.0 (32.0-70.0)	73.0 (41.0-92.0)	53.5 (31.2-76.0)	58.0 (38.8-83.8)
National ADI quartile, n/N (%)				
Least deprived ADI quartile (1-25)	36/55 (65.5)	0/55 (0.0)	10/55 (18.2)	10/55 (18.2)
Second least deprived ADI quartile (26-50)	41/70 (58.6)	4/70 (5.7)	13/70 (18.6)	19/70 (27.1)
Second most deprived ADI quartile (51-75)	52/87 (59.8)	3/87 (3.4)	17/87 (19.5)	15/87 (17.2)
Most deprived ADI quartile (76-100)	33/70 (47.1)	6/70 (8.6)	14/70 (20.0)	30/70 (42.9)

[a]Percentages may not add up to 100 due to rounding and missing values.
[b]eHEALS: eHealth Literacy Scale.

Regression Analysis of Barriers to EPP Use

In the overall regression including demographic variables, the ADI, and eHealth literacy level, only low eHealth literacy level (adjusted odds ratio [AOR] 1.32, 95% CI 1.13-1.54; *P*<.001) and older age (AOR 1.007, 95% CI 1.003-1.009; *P*<.001) predicted barriers to EPP access (Table 4). Similarly, only low eHealth literacy level and age were associated with predicting a technology discomfort–related barrier (low eHealth literacy, AOR 1.25, 95% CI 1.11-1.40; *P*<.001; age, AOR 1.004, 95% CI 1.002-1.007; *P*<.001; Table 4) or a preference-related barrier (low eHealth literacy, AOR 1.33, 95% CI 1.17-1.51; *P*<.001; age, AOR 1.004, 95% CI 1.00-1.01; *P*=.01; Table 4). The ADI was not associated with predicting overall barriers (*P*=.59), preference-related barriers (*P*=.35), or technology discomfort–related barriers (*P*=.76; Table 4).

Table . Regression of patient characteristics associated with self-reporting at least one barrier to electronic patient portal use (any barrier type, preference barrier, or technology discomfort barrier).

Characteristic	Regression 1: any barrier type[a]		Regression 2: preference barriers[b]		Regression 3: technology discomfort barriers[c]	
	AOR[d] (95% CI)	*P* value	AOR (95% CI)	*P* value	AOR (95% CI)	*P* value
Age	1.007 (1.003-1.009)	<.001	1.004 (1.00-1.01)	.01	1.004 (1.002-1.007)	<.001
Race or ethnicity						
Hispanic or Latino	0.99 (0.78-1.25)	.92	1.07 (0.88-1.31)	.49	1.04 (0.86-1.24)	.71
Non-Hispanic Black	1.09 (0.94-1.27)	.26	1.09 (0.95-1.24)	.21	1.04 (0.92-1.17)	.52
Non-Hispanic White	Reference	—[e]	Reference	—	Reference	—
Other identity or preferred not to answer	1.02 (0.79-1.32)	.88	1.05 (0.85-1.30)	.63	0.95 (0.78-1.15)	.56
National ADI[f]	1.00 (1.00-1.00)	.59	1.00 (1.00-1.00)	.35	1.00 (1.00-1.00)	.76
eHealth literacy level (eHEALS[g])						
High eHealth literacy	Reference	—	Reference	—	Reference	—
Low eHealth literacy	1.32 (1.13-1.54)	<.001	1.33 (1.17-1.51)	<.001	1.25 (1.11-1.40)	<.001
Subspecialty clinic						
Hand	Reference	—	Reference	—	—	—
Foot and ankle	1.01 (0.88-1.16)	.91	1.02 (0.91-1.15)	.73	—	—
Trauma	1.18 (0.99-1.41)	.07	1.17 (1.00-1.37)	.05	—	—
Highest educational level attained						
High school or lower	1.06 (0.90-1.25)	.65	—	—	—	—
Some college	0.96 (0.82-1.13)	.51	—	—	—	—
College or higher	Reference	—	—	—	—	—
Annual income bracket (US $)						
≤30,000	1.11 (0.94-1.30)	.22	—	—	—	—
30,001-50,000	1.05 (0.89-1.24)	.59	—	—	—	—
>50,000	Reference	—	—	—	—	—

[a]Regression model 1: 88.5% (254/287) of the patients were included after participants with missing values were excluded (outcome size: 124/254, 48.8% reported any barrier; df=12; pseudo-R^2=0.21).

[b]Regression model 2: 96.5% (277/287) of the patients were included after participants with missing values were excluded (outcome size: 78/277, 28.2% reported preference barriers; df=8; pseudo-R^2=0.16).

[c]Regression model 3: 96.5% (277/287) of the patients were included after participants with missing values were excluded (outcome size: 55/277, 19.9% reported technology discomfort barriers; df=6; pseudo-R^2=0.13).

[d]AOR: adjusted odds ratio.

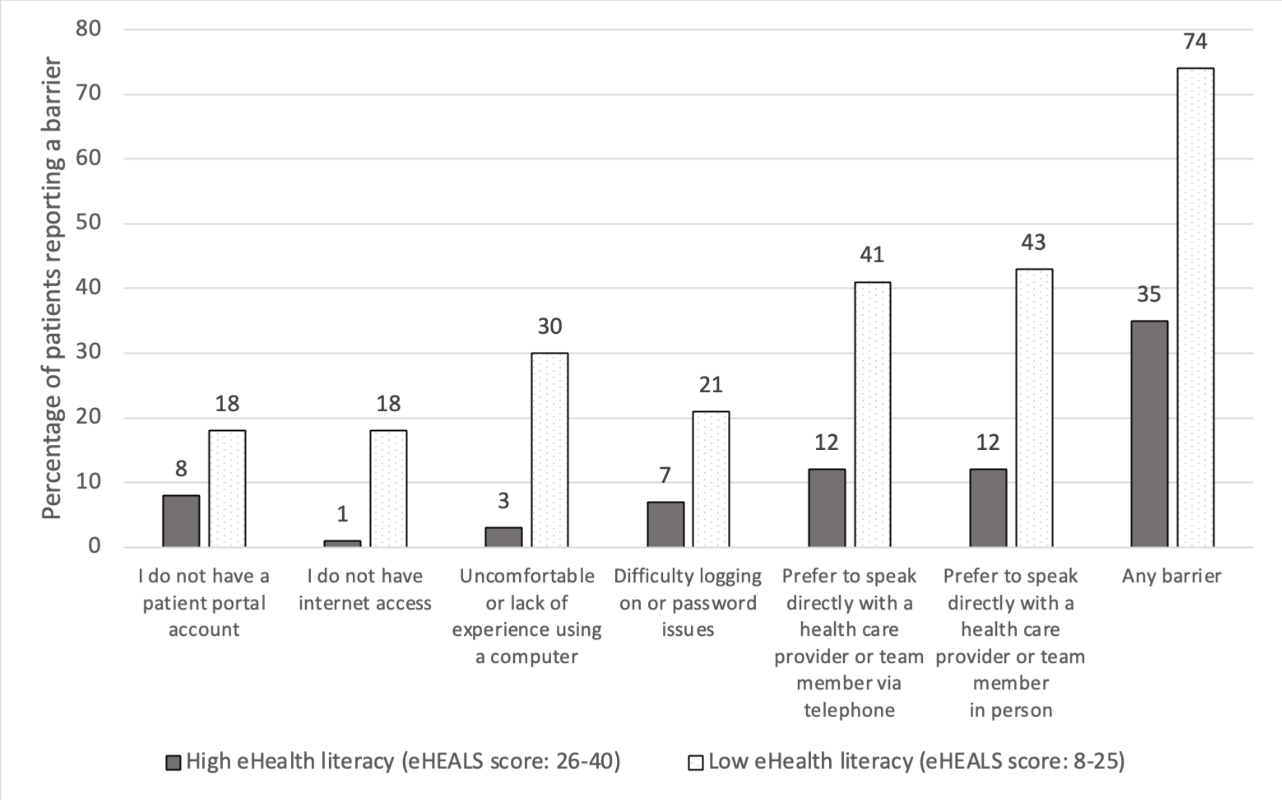
[e]Not applicable.

[f]ADI: Area Deprivation Index.

[g]eHEALS: eHealth Literacy Scale.

The most prevalent patient-reported barrier to EPP access was a preference to speak directly to the health care provider or team member in person or via telephone (52/287, 18.1%). A lack of comfort with a computer was cited as a barrier to EPP access by 29.5% (18/61) of patients with low eHealth literacy (Figure 1). Additionally, 17.8% (51/287) of study participants reported poor or fair ability to use a computer, tablet, or smartphone to find information that they needed on the internet. One or more barriers to EPP use were indicated by 80.4% (41/51) of patients with poor or fair self-ratings, compared to 35.3% (83/235) of patients with good, very good, or excellent self-ratings (*P*<.001).

Figure 1. Frequency and percentages of patient self-reported barriers to electronic patient portal (EPP) access stratified by eHealth literacy level. The percentages of patients who reported a particular barrier are reported over the denominator of patients with either high eHealth literacy scores (n=224) or low eHealth literacy scores (n=61). Percentages do not necessarily add up to 100 across barrier categories as patients could indicate more than one barrier. The barriers “I do not need a patient portal account,” “I have multiple patient portals,” and “I have privacy concerns” are not depicted separately due to lower frequency of observations.



Discussion

Principal Findings

Equitable implementation of digitized health tools relies on efforts from clinicians, researchers, and policymakers alike. This study assessed patient-reported barriers to use within an expanded framework of individual- and neighborhood-level factors. Low eHealth literacy level was the most significant determinant of overall barriers to EPP use for orthopedic surgery, followed by older age, as compared to other demographic factors and measures of neighborhood-level disadvantage. Contrary to expectation, neighborhood-level disadvantage as measured via the ADI had no mediating effect on barriers after adjusting for eHealth literacy level. Patients with physical access barriers did have appreciably higher ADIs; however, few patients reported physical access barriers overall. These findings build on prior work that showed that older age, among other patient demographics, was associated with reduced EPP enrollment in orthopedic surgery [3,4,8]. This also builds on prior work in the hospital medicine setting that showed that lower eHealth literacy correlated with decreased awareness and use and less favorable attitudes toward use of EPPs [13].

Barriers of physical access, such as lack of internet access, were infrequent compared to barriers related to preference or discomfort with technology, including lack of experience using a computer or difficulty logging in. This finding contrasts with previous work that found internet access to be a significant determinant of portal use for orthopedic surgery [6]. Additionally, among the 17.8% (51/287) of the participants with lower self-ratings of ability to use a computer, tablet, or smartphone to find information that they needed on the internet, 80.4% (41/51) indicated barriers to accessing their EPPs. The final rule of the 21st Century Cures Act alleviated many barriers related to physical access (eg, computer access or broadband coverage) by requiring an interoperability standard that any EPP programming interface be compatible with smartphone apps [29]. However, this legislation does not address barriers experienced by patients who may technically have the physical and digital tools to access their EPPs but not the self-efficacy or the skill sets to effectively use these tools.

These findings may support a second digital divide being dependent on disparities in skill sets rather than physical access [13,30]. With regard to EPPs, this is clinically significant as most institutions introduce EPPs via email with time-sensitive links and sign-up instructions. This may be ineffective at best in promoting EPP adoption or postenrollment use in populations who may have internet access yet lack the internet experience and skills to navigate setting up a digital account with time-pressure activation codes for protected health information. Moreover, the finding that over a quarter of all patients (52/287, 18.1%) perceived their preference to speak with a health care provider in person or via telephone as an actual barrier to using the EPP suggests that patients may view the EPP as a substitute for health care provider communication missing the personal element rather than as an adjunct to improve communication and transparency, as it was intended. This finding also suggests that it may be important to document patient preference for EPP versus non-EPP communication even if a patient does have an EPP as patients with activated EPPs may not be active users.

Importantly, while most patients were enrolled in the EPP, a significant portion of enrolled patients reported barriers to access and use of their portal. This is significant as prior studies of EPP use for orthopedic surgery have either only assessed EPP activation status as a surrogate for use and without assessing barriers to use [3,4,8] or qualitatively analyzed optional comments regarding nonuse in a small fraction of the study cohort (38 of 150 patients) [6]. These findings are clinically relevant as prior efforts to reduce disparities in EPP use have also focused on enrollment [9]. However, simply enrolling patients does little to address the underlying barriers of certain patients to using their EPPs. Enrollment will not address individual-level factors such as eHealth literacy, which may constitute a larger underlying barrier to sustained EPP use after enrollment (eg, patients' technological capabilities and skills rather than physical access to the technology itself).

Additionally, this study substantiates that older adults are a vulnerable population that may be left behind in a digitized health system. This is particularly critical to perioperative care in orthopedic surgery. Older patients may have more complex discharge needs, including perioperative medication changes and rehabilitation requiring close postoperative communications, and addressing them within the EPP may be ineffective. Older patients have intersecting factors that impede their access. While age and eHealth literacy were noncollinear in this study, age has well-known associations with traditional health and eHealth literacy [31]. Addressing deficiencies in these skill sets may improve lower levels of self-efficacy to adopt and use EPPs previously reported in older individuals [32]. This group may benefit from a proactive staff-level intervention that supports an in-person EPP enrollment option followed by an initial lesson on how to use the EPP, which Bhashyam et al [33] previously showed may be beneficial to improving postoperative follow-up. Notably, while older patients may have an elevated sense of caution in using online platforms due to counseling from groups such as the American Association of Retired Persons, patients infrequently noted privacy concerns as a barrier to EPP use in this study.

Patients with physical access barriers also had appreciably higher ADIs despite this analysis not meeting statistical significance. These patients may also benefit from routine touchpoints with staff outside of the EPP. Ensuring effective follow-up is especially important in these patients as, in addition to worse ADIs predicting worse comorbid chronic disease outcomes, simply living in a disadvantaged neighborhood confers similar readmission risk as having a chronic lung disease and higher risk than having a chronic condition such as diabetes [34]. At an informatics design level, a widget within the electronic medical record could be implemented to automatically yield an address-generated ADI analogously to how BMI is automatically calculated based on a patient's weight and height as this may generate similarly important contextual information to a patient's overall health, especially in a perioperative context.

Limitations

First, this was a single-institution study conducted across several orthopedic surgery subspecialty clinics not including adult reconstruction. Hence, the results may not be generalizable across other settings. Moreover, while eHEALS is a commonly used, validated screener for eHealth literacy in outpatient settings [18-20], it has not been validated in our specific patient population. Additionally, self-reporting via a survey may be limited by response bias; however, we felt that this method of examination was critical to include the patient perspective. Importantly, this study did not include non–English-speaking patients, who may experience additional barriers to EPP access; however, this study did include the perspectives of patients with limited reading and writing skills.

Additionally, level of use was dichotomized as routine and nonroutine, similar to the study by Maroney et al [17], without an option for "as needed," which assumes regular use. The National Cancer Institute's 2020 HINTS showed that only 40% of those with EPPs used it every year: 65% felt that they did not need to use it every year [11]. An additional checkbox option for "use as needed" may better capture this nuance; however, this could introduce indeterminate subjectivity.

Notably, the sample size may have been underpowered to detect small effect sizes in demographic differences and in novel outcomes such as the ADI. The sample size calculation was predicated on the 2020 HINTS, which analyzed barriers to EPP enrollment and did not include patients with activated EPPs who still experienced barriers (259/287, 90.2% of our study population) [11]. Moreover, no relevant existing literature has assessed the ADI. Finally, the secondary barrier categories were inductively coded by our research team after data collection based on natural groupings in which we were interested. This rendered our initial sample size calculation insufficient to perform a secondary regression analysis for the access-related barrier category, which had a small outcome size. Additionally, this classification was not validated, which may introduce bias but, importantly, allowed for discovery of new insights that may not have been generated through precoding.

Conclusions

Routine use of EPPs for online scheduling, patient communication, and telehealth continues to be a critical aspect of care. It is necessary to understand existing disparities in barriers to EPP access to not only improve access to care for all patients but also to continue building patients' toolbox and self-efficacy to take on active roles in their care. Future research should establish whether interventions, education, and improved eHealth literacy may overcome these barriers.

Acknowledgments

The authors would like to thank Jennifer Wolf, MD, and Megan Conti Mica, MD, for helping facilitate data collection within their outpatient clinics. The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content of this manuscript.

Funding

The authors did not receive support from any organization for the submitted work. No funding was received to assist with the preparation of this manuscript. No funding was received for conducting this study.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to institutional review board constraints but are available from the corresponding author on reasonable request.

Authors' Contributions

ALL, NL, KH, JS, and JGS contributed equally to the design, data collection, and writing of the manuscript.

Each author certifies that there are no funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article related to the author or any immediate family members.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of survey questions regarding patient portal.
[DOCX File, 15 KB - periop_v9i1e72035_app1.docx]

Multimedia Appendix 2
Summary of eHealth Literacy Scale items.
[DOCX File, 15 KB - periop_v9i1e72035_app2.docx]

References

1. Dendere R, Slade C, Burton-Jones A, Sullivan C, Staib A, Janda M. Patient portals facilitating engagement with inpatient electronic medical records: a systematic review. J Med Internet Res 2019 Apr 11;21(4):e12779. [doi: 10.2196/12779] [Medline: 30973347]
2. Davis Giardina T, Menon S, Parrish DE, Sittig DF, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. J Am Med Inform Assoc 2014;21(4):737-741. [doi: 10.1136/amiajnl-2013-002239] [Medline: 24154835]
3. Varady NH, d'Amonville S, Chen AF. Electronic patient portal use in orthopaedic surgery is associated with disparities, improved satisfaction, and lower no-show rates. J Bone Joint Surg Am 2020 Aug 5;102(15):1336-1343. [doi: 10.2106/JBJS.19.01080] [Medline: 32769600]
4. Vallurupalli N, Lawrence KW, Habibi AA, Bosco JA, Lajam CM. Socioeconomic disparities in online patient portal utilization among total knee arthroplasty recipients. J Arthroplasty 2024 Sep;39(9S2):S95-S99. [doi: 10.1016/j.arth.2024.04.040] [Medline: 38670173]
5. Plate JF, Ryan SP, Bergen MA, Hong CS, Attarian DE, Seyler TM. Utilization of an electronic patient portal following total joint arthroplasty does not decrease readmissions. J Arthroplasty 2019 Feb;34(2):211-214. [doi: 10.1016/j.arth.2018.11.002] [Medline: 30497899]
6. Kolodychuk NL, Wong M, Chimento G, Adams T, Gastanaduy M, Waddell BS. Web-based patient portal access in an orthopedic adult reconstruction patient population. Arthroplast Today 2019 Feb 22;5(1):83-87. [doi: 10.1016/j.artd.2019.01.004] [Medline: 31020029]
7. Fatehi A, Gonzalez A, Ring DC, Queralt M. Psychosocial factors are associated with electronic portal registration. Clin Orthop Relat Res 2020 Sep;478(9):2079-2084. [doi: 10.1097/CORR.0000000000001278] [Medline: 32332246]
8. Bell K, Warnick E, Nicholson K, et al. Patient adoption and utilization of a web-based and mobile-based portal for collecting outcomes after elective orthopedic surgery. Am J Med Qual 2018;33(6):649-656. [doi: 10.1177/1062860618765083] [Medline: 29562769]

9. Richwine C, Johnson C, Patel V. Disparities in patient portal access and the role of providers in encouraging access and use. J Am Med Inform Assoc 2023 Jan 18;30(2):308-317. [doi: 10.1093/jamia/ocac227] [Medline: 36451262]
10. Anthony DL, Campos-Castillo C, Lim PS. Who isn’t using patient portals and why? Evidence and implications from a national sample of US adults. Health Aff (Millwood) 2018 Dec;37(12):1948-1954. [doi: 10.1377/hlthaff.2018.05117] [Medline: 30633673]
11. Johnson C, Richwine C, Patel V. Individuals’ access and use of patient portals and smartphone health apps, 2020. In: ASTP Health IT Data Brief [Internet]: Office of the Assistant Secretary for Technology Policy; 2021.
12. Nishii A, Campos-Castillo C, Anthony D. Disparities in patient portal access by US adults before and during the COVID-19 pandemic. JAMIA Open 2022 Dec;5(4):ooac104. [doi: 10.1093/jamiaopen/ooac104] [Medline: 36540762]
13. Deshpande N, Arora VM, Vollbrecht H, Meltzer DO, Press V. eHealth literacy and patient portal use and attitudes: cross-sectional observational study. JMIR Hum Factors 2023 Jan 27;10:e40105. [doi: 10.2196/40105] [Medline: 36705947]
14. Lyles CR, Nelson EC, Frampton S, Dykes PC, Cemballi AG, Sarkar U. Using electronic health record portals to improve patient engagement: research priorities and best practices. Ann Intern Med 2020 Jun 2;172(11 Suppl):S123-S129. [doi: 10.7326/M19-0876] [Medline: 32479176]
15. Lyles CR, Nguyen OK, Khoong EC, Aguilera A, Sarkar U. Multilevel determinants of digital health equity: a literature synthesis to advance the field. Annu Rev Public Health 2023 Apr 3;44:383-405. [doi: 10.1146/annurev-publhealth-071521-023913] [Medline: 36525960]
16. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
17. Maroney K, Curtis LM, Opsasnick L, et al. eHealth literacy and web-based patient portal usage among kidney and liver transplant recipients. Clin Transplant 2021 Feb;35(2):e14184. [doi: 10.1111/ctr.14184] [Medline: 33278846]
18. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
19. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021 Nov 15;23(11):e30644. [doi: 10.2196/30644] [Medline: 34779781]
20. Cherid C, Baghdadli A, Wall M, et al. Current level of technology use, health and eHealth literacy in older Canadians with a recent fracture-a survey in orthopedic clinics. Osteoporos Int 2020 Jul;31(7):1333-1340. [doi: 10.1007/s00198-020-05359-3] [Medline: 32112118]
21. Richtering SS, Hyun K, Neubeck L, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4. [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
22. Vollbrecht H, Arora VM, Otero S, Carey KA, Meltzer DO, Press VG. Measuring eHealth literacy in urban hospitalized patients: implications for the post-COVID world. J Gen Intern Med 2021 Jan;36(1):251-253. [doi: 10.1007/s11606-020-06309-9] [Medline: 33078294]
23. Kind AJ, Buckingham WR. Making neighborhood-disadvantage metrics accessible - the Neighborhood Atlas. N Engl J Med 2018 Jun 28;378(26):2456-2458. [doi: 10.1056/NEJMp1802313] [Medline: 29949490]
24. The Python language reference. Python. URL: https://docs.python.org/3/reference/ [accessed 2022-08-31]
25. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar;17(3):261-272. [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
26. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. Presented at: Proceedings of the 9th Python in Science Conference; Jun 28 to Jul 3, 2010. [doi: 10.25080/Majora-92bf1922-011]
27. El-Toukhy S, Méndez A, Collins S, Pérez-Stable EJ. Barriers to patient portal access and use: evidence from the health information national trends survey. J Am Board Fam Med 2020;33(6):953-968. [doi: 10.3122/jabfm.2020.06.190402] [Medline: 33219074]
28. Lockwood MB, Dunn-Lopez K, Pauls H, Burke L, Shah SD, Saunders MA. If you build it, they may not come: modifiable barriers to patient portal use among pre- and post-kidney transplant patients. JAMIA Open 2018 Jul 10;1(2):255-264. [doi: 10.1093/jamiaopen/ooy024] [Medline: 31984337]
29. ONC’s Cures Act Final Rule. HealthIT.gov. URL: https://www.healthit.gov/topic/oncs-cures-act-final-rule [accessed 2025-12-24]
30. van Deursen A, van Dijk J. Internet skills and the digital divide. New Media Soc 2010;13(6):893-911. [doi: 10.1177/1461444810386774]
31. Chesser AK, Keene Woods N, Smothers K, Rogers N. Health literacy and older adults: a systematic review. Gerontol Geriatr Med 2016 Mar 15;2:2333721416630492. [doi: 10.1177/2333721416630492] [Medline: 28138488]
32. Price-Haywood EG, Harden-Barrios J, Ulep R, Luo Q. eHealth literacy: patient engagement in identifying strategies to encourage use of patient portals among older adults. Popul Health Manag 2017 Dec;20(6):486-494. [doi: 10.1089/pop.2016.0164] [Medline: 28384076]
33. Bhashyam AR, Bansal M, McGovern MM, van der Vliet QMJ, Heng M. In-hospital enrollment into an electronic patient portal results in improved follow-up after orthopedic surgery: cluster randomized controlled trial. JMIR Perioper Med 2022 Aug 11;5(1):e37148. [doi: 10.2196/37148] [Medline: 35969452]

34. Kind AJ, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014 Dec 2;161(11):765-774. [doi: 10.7326/M13-2946] [Medline: 25437404]

Abbreviations

ADI: Area Deprivation Index
AOR: adjusted odds ratio
eHEALS: eHealth Literacy Scale
EPP: electronic patient portal
HINTS: Health Information National Trends Survey
HIPAA: Health Insurance Portability and Accountability Act
REDCap: Research Electronic Data Capture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by N Rohatgi; submitted 03.Feb.2025; peer-reviewed by A Tsehay, M Rincon, X Ding; revised version received 06.Dec.2025; accepted 09.Dec.2025; published 07.Jan.2026.

Please cite as:
Litvak AL, Lin N, Hynes K, Strelzow J, Stepan JG
Assessing the Effects of eHealth Literacy and the Area Deprivation Index on Barriers to Electronic Patient Portal Use for Orthopedic Surgery: Cross-Sectional Observational Study
JMIR Perioper Med 2026;9:e72035
URL: https://periop.jmir.org/2026/1/e72035
doi:10.2196/72035

Benchmark Development for Fundamental Arthroscopic Skills Using a Simulation-Based Training Program: Observational Study

Eric Davis[1], MBA; Brianna Caraet[2], MD; Robert Pedowitz[3], MD, PhD; Gregg Nicandri[1], MD

[1]Department of Orthopaedic Surgery, University of Rochester, 601 Elmwood Avenue, Rochester, NY, United States

[2]Kaiser Permanente Orthopedic Sports Medicine, San Diego, CA, United States

[3]Department of Orthopaedic Surgery, University of California, Los Angeles, CA, United States

Corresponding Author:
Eric Davis, MBA
Department of Orthopaedic Surgery, University of Rochester, 601 Elmwood Avenue, Rochester, NY, United States

Abstract

Background: Surgical education has shifted from the traditional Halstedian apprenticeship model toward incorporating simulation due to work-hour restrictions, increasing case complexity, and economic and liability pressures. Building on the success of the Fundamentals of Laparoscopic Surgery program for general surgery, the Fundamentals of Arthroscopic Surgery Training (FAST) program was developed to establish proficiency benchmarks for orthopedic trainees in basic arthroscopic skills.

Objective: We aimed to establish benchmarks for 5 FAST workstation modules.

Methods: Sports medicine fellowship–trained faculty members were given instructions on the modules and 2 minutes of practice time, and they then performed each task 3 times with both their dominant and nondominant hand. For each module, mean faculty performance was used to establish an efficiency benchmark (time) and precision benchmark (errors).

Results: The Probing module should be completed in less than 95 seconds with no errors. The Ring Transfer module should be completed in less than 134 seconds with no more than 1 error. The Maze module should be completed in less than 99 seconds with no errors. The Meniscectomy module should be completed in less than 68 seconds with no more than 1 error. Lastly, the Suture Passing module should be completed in less than 195 seconds with no more than 1 error.

Conclusions: The FAST workstation can be used as a proficiency-based learning tool for residents to safely and effectively develop arthroscopic skills outside of the operating room. These benchmarks were established via a method previously validated in surgical simulation and balance precision and efficiency for skills that are considered generalizable and transferable to arthroscopic surgeries.

(JMIR Perioper Med 2026;9:e82723) doi:10.2196/82723

KEYWORDS

surgical education; orthopedic surgery; arthroscopy; simulation training; skills assessment; resident education

Introduction

Historically, surgical education has been based upon the Halstedian apprenticeship model, in which surgical trainees primarily acquire technical skills in the operating room while assisting and later performing procedures on live patients. In recent years, there has been a paradigm shift away from this approach in orthopedic surgery [1]. This can be attributed to multiple factors, including but not limited to resident work-hour restrictions, increased case complexity and variety, the risk of liability, and social and economic pressure. Simulation in surgical education has been suggested as a potential avenue to fill the gap in training, allowing trainees opportunities to obtain valuable experience in a low-risk, more cost-effective environment. In concordance with this, the Accreditation Council for Graduate Medical Education now requires that residency programs include surgical skills training in their curricula; this mandate has also been echoed by the American Board of Orthopaedic Surgery and the Residency Review Committee.

Simulation as an educational tool has been validated extensively in the field of general surgery. The Fundamentals of Laparoscopic Surgery program, developed in 1997 by the Society of American Gastrointestinal and Endoscopic Surgeons, combines cognitive and physical skills components shown to improve a given student's laparoscopic abilities over the course of the curriculum [2]. The program has been studied rigorously and validated since its release [3-9]. Currently, a passing score is required by the American Board of Surgery as prerequisite to sit for the certification exam. This has been well received by trainees, surgeons, and the public. As a result of this successful program in general surgery, the field of orthopedics has explored the development of a similar program.

In 2011, the Arthroscopy Association of North America (AANA), the American Board of Orthopaedic Surgery, and the American Academy of Orthopaedic Surgeons came together in a joint effort to establish the Fundamentals of Arthroscopic Surgery Training (FAST) program. The program is based on the tenet that fundamental surgical skills are best obtained sequentially, and that proficiency can be established through the successful completion of a basic skills curriculum. A curriculum and simulator have been developed for the FAST program and proficiency benchmarks for knot tying have been established.

The aim of this study was to develop time- and error-based proficiency benchmarks for 5 fundamental arthroscopic skills modules within a simulation-based training program.

Methods

Overview

The AANA offered a voluntary skills training course for residents and faculty at the Orthopaedic Learning Center in Chicago, Illinois, on December 7 to 9, 2017, and January 18 to 20, 2018. During downtime in the course, participants were recruited and provided with an information sheet outlining participation in this study. Two separate cohorts of 20 faculty each completed orthopedic simulations to improve their skill. For each module, the participant was given a list of written instructions and allowed to practice for a period of 2 minutes. Then, each participant performed the task 3 times with their dominant hand and 3 times with their nondominant hand. The order of dominant/nondominant completion was randomized so that exactly half of the faculty started with their nondominant hand. The time it took each participant to complete the module as well as the number of errors committed was recorded. Cohort 1 completed the Probing, Ring Transfer, and Maze Navigation modules. Cohort 2 completed the Meniscectomy and Suture Passing modules.

All modules were completed using camera-based visualization rather than live arthroscopes. Depending on the module, visualization was provided using either a USB camera system or a stationary camera, consistent with the simulation equipment available in the skills laboratory. This approach was used to standardize task execution across participants and to support reproducibility in a controlled training environment.

Five simulation-based arthroscopic skills modules were evaluated, including Probing, Ring Transfer, Maze Navigation, Partial Meniscectomy, and Suture Passage. Each module assessed fundamental psychomotor skills relevant to arthroscopy, with efficiency measured by task completion time and precision measured by predefined error criteria. Detailed descriptions of each module are provided in Multimedia Appendix 1.

Descriptive statistics were used for demographic variables, experience, handedness, efficiency (time) score, and precision (error) score. Mean performance time for all participants was determined. Any value lying more than 2 SDs from the mean was excluded. Exclusion of values greater than 2 SDs from the mean was performed to minimize the influence of extreme outliers that may reflect momentary lapses, equipment issues, or deviations from standardized task execution rather than true expert performance. This approach has been used in prior simulation-based benchmarking studies to derive stable proficiency thresholds that reflect typical expert performance rather than maximal or anomalous values. Outlier exclusion affected a small proportion of observations and did not meaningfully alter the relative ordering of module performance or the final benchmark thresholds. The trimmed mean was used as the basis for determining the proficiency level benchmark for each task. Efficiency benchmarks were set at the minimum threshold of the mean time of the expert cohort. Precision benchmarks were set at the mean number of errors committed by participants rounded to the nearest whole number. The benchmark was established as the mean performance when data from both hands were aggregated. Aggregation of dominant and nondominant hand performance was chosen a priori to reflect the ambidextrous demands of arthroscopic surgery, in which instrument laterality varies by procedure and portal placement. Use of the mean performance across both hands was intended to establish benchmarks that promote bilateral skill development while remaining feasible and generalizable for trainee progression.

Ethical Considerations

This study was reviewed by the University of Rochester Research Subjects Review Board and was granted exempt status with a waiver of documentation of informed consent. All participants were provided with an information sheet outlining the purpose, procedures, and voluntary nature of the study prior to participation. All data collected were deidentified. Each participant was assigned a unique study identification number that contained no personally identifiable information. No names or direct identifiers were recorded at any point in data collection or analysis. To protect participant privacy and confidentiality, data were recorded using study IDs only and stored securely in accordance with institutional research guidelines. Only members of the research team had access to the data. Results are presented in aggregate form to prevent identification of individual participants. Participation in this study was voluntary. Participants were not provided with any financial compensation or incentives for participation, and there were no costs associated with participation.

Results

Demographics

A total of 40 sports medicine fellowship–trained faculty surgeons were included in the analysis. The mean age of the participants was 43.9 (SD 9.0) years. They had an average of 10.5 (SD 8.3) years in practice and performed approximately 277 (SD 125) arthroscopic cases per year. All were male and 3 (8%) were left-handed. There were no significant differences in demographic data between cohorts from the 2 courses (Table 1).

Table . Demographics.

	Cohort 1 (n=20)	Cohort 2 (n=20)	Overall (n=40)	*P* value[b]
Age, mean (SD)	43.1 (8.38)	44.7 (9.22)	43.9 (9.0)	.59
Male gender, n (%)	20 (100)	20 (100)	40 (100)	—[a]
Percent fellowship trained, n (%)	20 (100)	20 (100)	40 (100)	—
Years in practice, mean (SD)	9.5 (7.7)	11.4 (8.5)	10.5 (8.3)	.48
Cases per year (n), mean (SD)	275 (130.1)	279 (117.3)	277 (125)	.94
Right-handed, n (%)	18 (90)	19 (95)	37 (93)	.55

[a]Not applicable.

[b]For age, number of years in practice, and number of cases per year, 2-sided *t* tests were used to test for statistical significance. For handedness, the χ^2 test was used. No differences were statistically significant.

Probing

For the Probing module, mean completion times were 92.5 (SD 17.6) seconds for the dominant hand and 97.3 (SD 13.6) seconds for the nondominant hand, with no errors recorded and no significant difference between hands (*P*=.35); the benchmark was <95 seconds with no errors.

Ring Transfer

For the Ring Transfer module, mean completion times were 117.5 (SD 26.7) seconds for the dominant hand and 151.3 (SD 38.1) seconds for the nondominant hand, with mean error rates of 0.6 (SD 0.8) and 1.1 (SD 0.9) dropped rings, respectively; dominant and nondominant hand performance differed significantly (*P*<.001), and the benchmark was <134 seconds with no more than 1 dropped ring.

Maze

For the Maze module, mean completion times were 89.1 (SD 24.3) seconds for the dominant hand and 108.6 (SD 25.0) seconds for the nondominant hand, with mean error rates of 0.2 (SD 0.4) and 0.1 (SD 0.3) lost balls, respectively; dominant and nondominant hand performance differed significantly (*P*=.02), and the benchmark was <99 seconds with no errors.

Meniscectomy

For the Meniscectomy module, mean completion times were 73.2 (SD 20.7) seconds for the dominant hand and 63.7 (SD 23.6) seconds for the nondominant hand, with mean error rates of 0.1 (SD 0.3) over- or underresection events for both hands; no significant difference was observed between dominant and nondominant hand performance (*P*=.14), and the benchmark was <68 seconds with no more than 1 area of over- or underresection.

Suture Passing

For the Suture Passing module, mean completion times were 177.5 (SD 41.2) seconds for the dominant hand and 211.9 (SD 40.9) seconds for the nondominant hand, with mean error distances of 0.5 (SD 0.6) mm and 0.6 (SD 0.8) mm from the target, respectively; dominant and nondominant hand performance differed significantly (*P*<.001), and the benchmark was <195 seconds with no more than 1 mm from the target and no suture anchor unload.

For all modules, benchmark time was defined as the mean performance across dominant and nondominant hands to reflect bilateral task demands. For example, if the right hand took 60 seconds and the left hand took 30 seconds, the benchmark time was 45 seconds. Benchmarks for all modules are summarized in Table 2.

Table . Benchmarks.

Module	Benchmark time (seconds)[a]	Benchmark errors
Probing module	<95	No errors
Ring Transfer module	<134	No more than 1 dropped ring
Maze module	<99	No balls off platform
Meniscectomy module	<68	Only 1 area of over- or underresection
Suture Passing module	<195	<1 mm from the target area and no suture anchor unloads

[a]Benchmark time is the mean performance for both hands. For example, if the right hand takes 60 seconds, and the left hand takes 30 seconds, the time to compare against the benchmark is 45 (90/2) seconds.

Discussion

Principal Findings

In this study, we established objective time- and error-based proficiency benchmarks for 5 fundamental arthroscopic skills modules using performance data from fellowship-trained sports medicine faculty members. These benchmarks were derived using a standardized, previously validated simulation-based methodology and were designed to balance efficiency and precision across tasks of increasing technical complexity. Defining proficiency thresholds based on expert performance involves an inherent trade-off. Benchmarks that are too lenient may allow progression without adequate skill consolidation, whereas overly stringent thresholds may discourage practice or be impractical within residency training constraints. By deriving benchmarks from aggregated expert performance and excluding extreme outliers, the present thresholds were designed to balance feasibility with meaningful skill assessment. Continued evaluation of trainee outcomes and future transfer-validity studies will be important to further calibrate these benchmarks.

The FAST program consists of a curriculum, a simulator, and a series of tests, which enable trainees to iteratively practice fundamental skills outside of the operating room and ideally before performing surgery on patients. Training such motor skills outside of the OR is crucial in arthroscopy, which has a steep learning curve [10]. The FAST program's hands-on design is in alignment with the Institute of Medicine's recommendations for graduate medical education, which aim to transition from process-driven training to a more outcome-driven approach. Instead of the assumption that residents acquire their needed skills after completing a certain number of procedures or after a certain amount of time in training, the institute recommends evaluating proficiency in such skills before advancement [11]. The FAST program uses this proficiency-based progression approach to training to encourage "deliberate practice" to improve motor skills [12]. With the establishment of benchmarks for each FAST module, a resident is able to receive real-time objective feedback on their performance and is encouraged to strive to be comparable to experienced arthroscopy surgeons. With the accessibility of the workstations, residents are able to obtain additional practice repetitions and adjust their technique as needed to meet the established standard. Recent evaluations of the FAST workstation have demonstrated measurable differences in performance across postgraduate training levels and institutions, further supporting the need for standardized, objective benchmarks [13].

The FAST program can be integrated as early as the intern year in orthopedic training so that young surgeons can develop foundational psychomotor skills before even touching an arthroscope in the operating room. This early exposure to arthroscopic instruments and techniques can instill confidence in a new resident and their abilities, help them develop good habits early, and potentially decrease patient morbidity. Studies have demonstrated that medical students and residents who train with simulators perform better in the operating room than their counterparts without simulator training, with recent arthroscopy-specific work demonstrating transfer validity to diagnostic knee arthroscopy and meniscectomy performance [14,15]. With improved resident performance in the operating room, a faculty member walking a resident through the case would likely feel more comfortable increasing resident autonomy, and valuable operating room time can be spent teaching and instructing on the unique nuances of a particular case instead of on basic skills.

The FAST workstations are easily accessible, cost-effective, and efficient. They require a simple set-up, and at our institution, they are available in the skills laboratory for all residents to use at any time. Residents can easily independently set up the workstation in the laboratory as long as they have an arthroscopy tower or USB camera with which to practice. Currently, we have each of the testing modules set up in front of a computer monitor with all the instruments available that are necessary to complete the task. We have a QR code that links to a video of the task and instructions mounted on the wall above each of the modules (Figure 1). We use iPevo cameras when a stationary light source is sufficient (Knot Tying) and Sawbones USB cameras when a replica of an arthroscope is required (Ring Transfer). We have found that having easy-to-access instructions and all modules already set up increases use of the simulator and maximizes resident time. We typically introduce FAST during a 2-to-3-hour session with all interns that is proctored by an instructor during their skills month. This familiarizes them with the simulator and modules. They are expected to practice to proficiency and then use the simulators to refine their skills prior to their arthroscopic rotations. The modular design of the FAST program also supports scalability across institutions with varying resource availability and learner profiles, as the workstations can be implemented using differing levels of visualization technology and integrated at multiple stages of training. This flexibility allows programs to adapt FAST implementation to local constraints while maintaining standardized proficiency benchmarks.

Figure 1. Simulation laboratory set-up.

Prior papers have aimed to establish clear quantitative benchmarks for simulation training based on the mean performance of an experienced surgeon cohort [16]. One such paper, by Pedowitz et al [17], established the benchmark for the Knot Tying module in the FAST program, which is the last module to be completed in the series. It evaluated the performance of 50 faculty members in attendance at AANA resident arthroscopy courses. The 2 faculty cohorts had an average of 19.3 and 19.9 years of experience in practice. A benchmark was established using data from the more successful of the 2 faculty cohorts and established proficiency as equal to 2 knot failures or less out of 5 knot attempts when using the knot tester workstation. This randomized, prospective study then evaluated a training group of 44 postgraduate year 4 or 5 orthopedic residents. They were divided into 3 subgroups: group A received standard didactic training, group B was allowed additional knot-tying workstation practice, and group C received proficiency-based progression training with the knot-tying workstation. While the aggregate resident knot failure rate of 26% was higher than the 22% knot failure rate of the faculty, resident group C had only an 11% knot failure rate with 94% of residents in this group passing the threshold. This paper both demonstrated a system to establish a benchmark standard and proved that proficiency-based progression training improves the likelihood of meeting said benchmarks. Recent FAST-based educational interventions have demonstrated improvements in objective performance metrics, including arthroscopic knot integrity, when guided, proficiency-based training is used [18]. The arthroscopic knot-tying station is the last in the sequence of the FAST workstation modules as it tests one of the more complex skills in arthroscopy. The purpose of the current investigation was to establish objective benchmarks for the remaining 5 FAST workstation modules: Probing, Ring Transfer, Maze Navigation, Meniscectomy, and Suture Passing.

These benchmarks derived from the performance of fellowship-trained sports medicine faculty members were designed to be both realistic and achievable for residents. The 6 modules progress from basic skills to more complex tasks. The first module, Probing, allows residents to establish basic arthroscopic skills of horizontal control, telescoping, periscoping, and triangulation in order to feel comfortable performing a diagnostic arthroscopy. Gaining familiarity with the arthroscope in this module lays the foundation for establishing visualization in the remainder of the tasks and in the operating room. The second module, Ring Transfer, reinforces these skills and allows the resident to develop proficiency with the arthroscopic grasper, which is used when they are tasked with performing an arthroscopic loose-body removal. The dexterity with the instruments developed in this module is applicable to using arthroscopic instruments to manipulate tissues or implants in all arthroscopic procedures. The next module, Maze Navigation, builds on prior probing skills from the first module and develops tracking skills with nonstationary objects. This skill is necessary to develop before interventional arthroscopy can be performed. While the first 3 modules focus more on skills generalization, the final 3 modules correlate even more closely with specific surgical interventions, thereby focusing more on skills transfer. The fourth module directly simulates completing a partial meniscectomy with a benchmark designed to encourage an adequate amount of resection. During this practice, residents develop more familiarity with a biter and the ability to maneuver more deliberately and precisely. Resection of a paper "meniscus" is a safer alternative to a new resident damaging cartilage with the biter or over-resecting and destabilizing a meniscus tear because they are using a biter for the first time in an actual patient's knee. The second to last task, Suture Passing, integrates the use of other new instruments, including an antegrade suture passer, a piercing suture passer/retriever, and a suture lasso or suture shuttle. This task's surgical correlate is an arthroscopic rotator cuff repair or Bankart repair, which are among the more complex arthroscopic procedures. Completion of this task relies on the mastery and coordination of skills obtained in the previous modules. The last module, Knot Tying, allows residents to evaluate the biomechanical integrity of their arthroscopic knots in a setting safer than discovering loose knots in a failed arthroscopic suture repair. In conclusion, a resident's ability to meet the time constraints and limit their errors for each task translates into more efficient and accurate performance in the operating room, which promotes patient safety.

In establishing the benchmarks, there was a statistically significant difference in performance between the faculty surgeons' dominant and nondominant hands with the Ring Transfer, Maze, and Suture Passing modules. However, surgeons have to be ambidextrous. Which hand holds the arthroscope and which holds the remaining instruments is often based on the laterality of the procedure. Previous studies have shown that experienced arthroscopic surgeons are more ambidextrous than novices [19]. Specifically, experts demonstrate significantly smaller dominant-nondominant differences in task completion time and error rates on simulated arthroscopic tasks compared with novices, supporting bilateral performance assessment as a marker of skill acquisition. Therefore, the FAST program requires all tasks to be completed with the right hand and the left hand in separate attempts in order to develop skills in both hands regardless of hand dominance. Establishing separate benchmarks for the dominant and the nondominant hand based on the differences seen in the expert cohorts was considered. Aggregated performance metrics are commonly used in established simulation benchmarking frameworks, such as the Objective Structured Assessment of Technical Skill (OSATS) and the Fundamentals of Laparoscopic Surgery program, which define proficiency using composite measures rather than limb-specific performance. This supports the use of bilateral performance means when establishing reproducible and generalizable training benchmarks. That said, the scientific ideal needs to be weighed equally with practicality in this case. The FAST Program puts multiple new demands on trainees and faculty in the programs that adopt it. Ultimately, in order to be successfully adopted, it needs to be easy to implement. Based upon our experience facilitating many FAST courses at the Orthopaedic Learning Center in Chicago and at multiple residency programs, we determined that a single benchmark for each task was necessary.

We also discussed whether to take the mean performance of both hands or to set the benchmark at the level of the dominant hand. We chose the mean performance. This was the best compromise between ensuring adequate ambidexterity while also recognizing that differences exist, even in expert faculty members. The FAST program is designed to enhance surgical skills and promote deliberate, repetitive practice for benchmarks. We were concerned that setting the benchmark as the performance of the dominant hand of the faculty members would have made it significantly harder to achieve for the trainees. The goal was for residents to be able to pass the FAST program by the end of their third year in training, and we decided that the likely significant increase in repetitions required to meet the more stringent benchmark would be frustrating and time consuming for residents who already have many other competing demands. Passing the FAST program at the end of the third year does not mean that residents are proficient at arthroscopy. It means that they have sufficiently practiced and likely have improved their fundamental arthroscopic skills by meeting a minimum threshold of performance. They still have additional years in residency to build upon the skills acquired in the FAST program. Future validation of FAST benchmarks could be pursued through prospective randomized or longitudinal studies comparing proficiency-based FAST training with standard curricula. Studies should be powered to detect differences in operative performance and stratified by postgraduate year to assess differential benefit by training stage. Translation to operating room performance may be measured using objective metrics such as task-specific error rates, time to task completion, and validated global rating scales (eg, OSATS) during comparable arthroscopic procedures. This benchmarking is foundational work that is required to enable such an investigation.

The development of the FAST program benchmarks is one of the factors that allow this to be a self-sufficient progression learning tool, setting it apart from other skills programs. In 2007, the Carolinas Medical Center integrated a laparoscopic skills curriculum into their general surgery residency program [20]. A skills laboratory coordinator needed to be present to record the duration of participation in each task and keep track of errors. In contrast, the FAST program is designed as a web-based program residents can access independently at a time convenient for them. A stopwatch is built into the program to record time to complete a task, and each task has a section to input the number of errors. These factors determine progression to the next module, all of which can be done at the resident's own pace. The paper about the Carolinas Medical Center program by Stefanidis et al [20] describes feedback occurring after each training session or when a resident is called in by the coordinator. With the FAST program, instant feedback is available with the clear benchmarks established for each task.

Limitations and Future Studies

A limitation of this study was that the included participants were exclusively faculty members teaching at the AANA Fundamentals of Arthroscopic Surgery Residents Course. All participants were male, sports medicine fellowship–trained faculty members, which may limit generalizability across genders, subspecialties, and training environments. Ultimately, this was a convenience sample due to the significant logistical hurdles to obtaining data with another method from a large number of expert orthopedic surgeons. This may have introduced a selection bias for surgeons, and they may not be fully representative of a wider population of sports fellowship–trained surgeons. The included cohort were deemed experts by virtue of their completed training, but there were outliers in performance among the participants. There were 21 (of 320) outlier data points, and these were subsequently excluded from the dataset before formulating the benchmarks.

Another limitation of this study is that it did not demonstrate the feasibility of the FAST program or its transfer validity, necessitating further studies. Further investigation is also needed to determine the average number of repetitions and practice time required to meet the proficiency benchmarks. These results could then be subanalyzed based on training level. This would provide a clearer estimated duration of FAST programs for other orthopedic residency programs interested in integrating the FAST program into their curricula. Further application of the FAST program could also be used to assess whether those who reach proficiency benchmarks perform better than those with no simulation training on other simulated surgical procedures and in the operating room for comparable surgical procedures. Lastly, further studies are needed to investigate which subset of residents achieve the most benefit from the FAST program as determined by the greatest improvement in arthroscopic skills. This could be used to determine in which year of residency training the FAST program should be integrated.

Benchmarks in this study were derived on a module-specific basis using different expert cohorts and should be interpreted as reference values for individual skills rather than curriculum-level proficiency standards. This design limits direct cross-module comparison but reflects practical constraints of expert data collection and supports initial benchmark development for discrete arthroscopic tasks. Evaluation of learners across the full FAST curriculum will be necessary to assess proficiency progress and program-level validity.

Conclusions

FAST workstations can be used as proficiency-based learning tools for residents to safely and effectively develop arthroscopic skills outside of the operating room. These benchmarks were established via a method previously validated in surgical simulation and balance precision and efficiency for skills that are considered generalizable and transferable to arthroscopic surgeries.

Funding

This study was supported by an educational grant from the Arthroscopy Association of North America (AANA) in the amount of $44,000. Study activities were conducted in part at the AANA Orthopaedic Learning Center. The funding source had no role in the design of the study; analysis; interpretation of the data; or in the writing of the manuscript or decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed procedural descriptions.
[DOCX File, 2779 KB - periop_v9i1e82723_app1.docx]

References

1. Pedowitz RA, Marsh JL. Motor skills training in orthopaedic surgery: a paradigm shift toward a simulation-based educational curriculum. J Am Acad Orthop Surg 2012 Jul;20(7):407-409. [doi: 10.5435/JAAOS-20-07-407] [Medline: 22751159]
2. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. Am J Surg 1998 Jun;175(6):482-487. [doi: 10.1016/s0002-9610(98)00080-4] [Medline: 9645777]
3. McCluney AL, Vassiliou MC, Kaneva PA, et al. FLS simulator performance predicts intraoperative laparoscopic skill. Surg Endosc 2007 Nov;21(11):1991-1995. [doi: 10.1007/s00464-007-9451-1] [Medline: 17593434]
4. Soper NJ, Fried GM. The fundamentals of laparoscopic surgery: its time has come. Bull Am Coll Surg 2008 Sep;93(9):30-32. [Medline: 19469354]
5. Scott DJ, Bergen PC, Rege RV, et al. Laparoscopic training on bench models: better and more cost effective than operating room experience? J Am Coll Surg 2000 Sep;191(3):272-283. [doi: 10.1016/s1072-7515(00)00339-2] [Medline: 10989902]
6. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002 Oct;236(4):458-463. [doi: 10.1097/00000658-200210000-00008] [Medline: 12368674]
7. Hyltander A, Liljegren E, Rhodin PH, Lönroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. Surg Endosc 2002 Sep;16(9):1324-1328. [doi: 10.1007/s00464-001-9184-5] [Medline: 11988802]
8. Korndorffer JR Jr, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ. Simulator training for laparoscopic suturing using performance goals translates to the operating room. J Am Coll Surg 2005 Jul;201(1):23-29. [doi: 10.1016/j.jamcollsurg.2005.02.021] [Medline: 15978440]
9. Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. Am J Surg 2010 Jan;199(1):115-120. [doi: 10.1016/j.amjsurg.2009.07.035] [Medline: 20103076]
10. Atesok K, Mabrey JD, Jazrawi LM, Egol KA. Surgical simulation in orthopaedic skills training. J Am Acad Orthop Surg 2012 Jul;20(7):410-422. [doi: 10.5435/JAAOS-20-07-410] [Medline: 22751160]
11. Committee on the Governance and Financing of Graduate Medical Education, Board on Health Care Services, Institute of Medicine. In: Eden J, Berwick DM, Wilensky GR, editors. Graduate Medical Education That Meets the Nation's Health Needs: National Academies Press (US); 2014. [Medline: 25340242]
12. Thomas GW, Johns BD, Marsh JL, Anderson DD. A review of the role of simulation in developing and assessing orthopaedic surgical skills. Iowa Orthop J 2014;34:181-189. [Medline: 25328480]

13. Chapek M, Otlans PT, Buuck T, et al. Resident performance on the Fundamentals of Arthroscopic Surgery Training workstation does not predictably improve with postgraduate year. Arthrosc Sports Med Rehabil 2024 Feb;6(1):100866. [doi: 10.1016/j.asmr.2023.100866] [Medline: 38318395]
14. Palter VN, Grantcharov T, Harvey A, Macrae HM. Ex vivo technical skills training transfers to the operating room and enhances cognitive learning: a randomized controlled trial. Ann Surg 2011 May;253(5):886-889. [doi: 10.1097/SLA.0b013e31821263ec] [Medline: 21394017]
15. Tronchot A, Casy T, Vallee N, et al. Virtual reality simulation training improve diagnostic knee arthroscopy and meniscectomy skills: a prospective transfer validity study. J Exp Orthop 2023 Dec 14;10(1):138. [doi: 10.1186/s40634-023-00688-8] [Medline: 38095746]
16. Gallagher AG, O'Sullivan GC. Fundamentals of Surgical Simulation: Principles and Practice: Springer; 2011.
17. Pedowitz RA, Nicandri GT, Angelo RL, Ryu RKN, Gallagher AG. Objective assessment of knot-tying proficiency with the Fundamentals of Arthroscopic Surgery Training program workstation and knot tester. Arthroscopy 2015 Oct;31(10):1872-1879. [doi: 10.1016/j.arthro.2015.06.021] [Medline: 26298642]
18. Farooq H, Gaetano A, Salazar D, Garbis N. Guided education and the Fundamentals of Arthroscopy Surgery Training (FAST) workstation improve surgical resident knot-tying skills. Arthrosc Sports Med Rehabil 2025 Jun;7(3):101110. [doi: 10.1016/j.asmr.2025.101110] [Medline: 40692921]
19. Pedowitz R, Nicandri G, Tuchschmid S. Asymmetry in dominant / non-dominant hand performance differentiates novices from experts on an arthroscopy virtual reality serious game. Stud Health Technol Inform 2016;220:289-294. [Medline: 27046594]
20. Stefanidis D, Acker CE, Swiderski D, Heniford BT, Greene FL. Challenges during the implementation of a laparoscopic skills curriculum in a busy general surgery residency program. J Surg Educ 2008;65(1):4-7. [doi: 10.1016/j.jsurg.2007.11.009] [Medline: 18308276]

Abbreviations

AANA: Arthroscopy Association of North America
FAST: Fundamentals of Arthroscopic Surgery Training
OSATS: Objective Structured Assessment of Technical Skill

Edited by N Rohatgi; submitted 25.Aug.2025; peer-reviewed by D Demirel, W Li; revised version received 22.Feb.2026; accepted 25.Feb.2026; published 08.Apr.2026.

Please cite as:
Davis E, Caraet B, Pedowitz R, Nicandri G
Benchmark Development for Fundamental Arthroscopic Skills Using a Simulation-Based Training Program: Observational Study
JMIR Perioper Med 2026;9:e82723
URL: https://periop.jmir.org/2026/1/e82723
doi:10.2196/82723

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