

**Psychometric assessment of PROMIS-29 as a measure of recovery after
colorectal surgery.**

Francesca Fermi, MD

Department of Experimental Surgery

McGill University, Montreal

November 2024

A thesis submitted to McGill University in partial fulfilment of the
requirements of the degree of Master of Science.

© Fermi Francesca, 2024

TABLE OF CONTENTS

ABSTRACT	2
RÉSUMÉ.....	3
STATEMENT OF ORIGINALITY.....	5
ACKNOWLEDGEMENTS.....	6
AUTHORS' CONTRIBUTIONS.....	7
LIST OF TABLES & FIGURES	8
LIST OF ABBREVIATIONS.....	9
CHAPTER 1 -INTRODUCTION	11
1.1 Recovery After Colorectal Surgery	11
1.2 Perioperative Care in Colorectal Surgery	12
1.3 Measurement of Postoperative Recovery	13
1.3.1 Traditional Measures	13
1.3.2 Patient-reported Outcomes Measures (PROMs)	14
1.3.3 Psychometric Properties of PROMs	16
1.3.4 Patient-Reported Outcome Measurement Information System (PROMIS)	18
1.4 Research gap and study aim.....	20
CHAPTER 2 -MANUSCRIPT	21
2.1 Introduction.....	22
2.2 Material and Methods	23
2.3 Results.....	28
2.4 Discussion.....	35
2.5 Conclusion.....	38
2.6 Disclosures.....	39
CHAPTER 3 - CONCLUSION AND FUTURE DIRECTIONS.....	40
REFERENCES	42
APPENDIX	58

ABSTRACT

Background: The PROMIS-29 questionnaire assesses general physical and mental health aspects that may be relevant to surgical recovery. However, evidence regarding the psychometric properties PROMIS-29 in the context of colorectal surgery remains limited. This study aims to assess the content validity, internal consistency, construct validity, and responsiveness of PROMIS-29 as a measure of recovery after colorectal surgery.

Methods: A cohort of adult patients (≥ 18 years) undergoing elective colorectal resection at two academic hospitals completed the PROMIS-29 questionnaire preoperatively and weekly until the postoperative week (POW) 4. Physical Health Summary (PHS) and Mental Health Summary (MHS) scores were derived from PROMIS-29 t-scores (higher scores indicating better health status). Content validity was assessed by comparing PROMIS-29 items against the conceptual framework of recovery after abdominal surgery. Internal consistency was assessed using Cronbach's alpha. Construct validity and responsiveness were assessed by testing *a priori* hypotheses.

Results: Overall, 282 patients were included (59 ± 15 years, 48% female, 78% minimally invasive, median hospital stay 3 days). PROMIS-29 items covered 36% of health domains considered important by patients. Internal consistency was acceptable (≥ 0.7). The construct validity of PHS and MHS scores was limited (0 out of 4 hypotheses confirmed on POW1; 1 or 2 hypotheses confirmed on POW2-4). Responsiveness was supported before POW2, but not after.

Conclusion: This study supports that PROMIS-29 has limited content validity, construct validity, and responsiveness when used to measure recovery within 4 weeks after colorectal surgery. There remains a need to develop psychometrically sound PROMs to measure postoperative recovery.

RÉSUMÉ

Introduction : Le questionnaire PROMIS-29 évalue les aspects généraux de la santé physique et mentale qui peuvent être pertinents pour la récupération chirurgicale. Cependant, les preuves concernant les propriétés psychométriques du PROMIS-29 dans le contexte de la chirurgie colorectale restent limitées. Cette étude vise à évaluer la validité du contenu, la cohérence interne, la validité de construit et la réactivité du PROMIS-29 en tant que mesure de la récupération après une chirurgie colorectale.

Méthodes : Une cohorte de patients adultes (≥ 18 ans) subissant une résection colorectale élective dans deux hôpitaux universitaires a complété le questionnaire PROMIS-29 en période préopératoire, puis chaque semaine jusqu'à la 4^e semaine postopératoire (POW4). Les scores de la santé physique (PHS) et de la santé mentale (MHS) ont été dérivés des scores t du PROMIS-29 (des scores plus élevés indiquant un meilleur état de santé). La validité du contenu a été évaluée en comparant les items du PROMIS-29 au cadre conceptuel de la récupération après une chirurgie abdominale. La cohérence interne a été évaluée à l'aide de l'alpha de Cronbach. La validité de construit et la réactivité ont été évaluées via des tests d'hypothèses de groupes connus.

Résultats : Au total, 282 patients ont été inclus (59 ± 15 ans, 48 % de femmes, 78 % en chirurgie mini-invasive, séjour hospitalier médian de 3 jours). Les items du PROMIS-29 couvraient 36 % des domaines de santé jugés importants par les patients. La cohérence interne était acceptable (>0.7). La validité de construit des scores PHS et MHS était limitée (aucune hypothèse confirmée à POW1; 1 ou 2 hypothèses confirmées à POW2-4). La réactivité était confirmée avant POW2, mais pas après.

Conclusion : Cette étude montre que le PROMIS-29 a une validité de contenu, une validité de construit et une réactivité limitées lorsqu'il est utilisé pour mesurer la récupération au cours des 4 semaines suivant une chirurgie colorectale. Il reste nécessaire de développer des

instruments de mesure des résultats psychométriquement solides pour évaluer la récupération postopératoire.

STATEMENT OF ORIGINALITY

The research outlined in this thesis constitutes an original contribution to the existing literature on the use of patient-reported outcome measures in perioperative care. Although I have benefited from the guidance of my supervisor, Research Advisory Committee members, and manuscript co-authors, the content presented in the subsequent chapters is entirely my own original work.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to everyone who contributed to the completion of this thesis.

First, my deepest thanks go to my supervisor, Julio F. Fiore Jr., who guided me through this project with boundless patience. He shared his expertise in methodology, statistics, and scientific writing, making it a true privilege to work alongside him. His mentorship has left a lasting impact on me, and I hope that the connection we've forged will continue to support me throughout my career.

I am grateful to the members of my Research Advisory Committee and my Chair for their essential guidance and insightful suggestions that helped bring this master's thesis to completion. I also want to thank my lab colleagues, who have become like family to me over this past year, especially Mak Pook, Katy Dmowski, and Joao de Cerqueira. A special thank you goes to Samin Shirzadi, whose dedication and expertise have been invaluable assets at work. But beyond her professional support, she has become a truly cherished friend. Meeting her has been a privilege, adding depth and meaning to this entire experience.

I also extend my gratitude to the Pancreas and Transplant Surgery team at San Raffaele Hospital in Milan. A special thanks to Nicolò Pecorelli, Stefano Partelli, Domenico Tamburrino, and Professor Massimo Falconi for their belief in me from the very first day and for allowing me to embrace this extraordinary opportunity.

Lastly, heartfelt gratitude is extended to friends and family for their unwavering support throughout this year, despite the ocean that separated us.

AUTHORS' CONTRIBUTIONS

Francesca Fermi (Primary author): study design, project coordination, data collection, statistical analysis, data interpretation, preparation of manuscript and thesis.

Samin Shirzadi: study design, project coordination, data collection, and review of the manuscript and thesis.

Ghadeer Olleik: study design, participant recruitment, data collection, review of the manuscript.

Makena Pook: study design, data collection, and review of the manuscript.

Maxime Lapointe Gagner: study design, data collection, and review of the manuscript.

Sarah Al Ben Ali: study design, data collection, and review of the manuscript.

Philip Nguyen Powanda: study design, data collection, and review of the manuscript.

Elahe Khorasani: study design, data collection, and review of the manuscript.

Tahereh Najafi Ghazeljeh: study design, data collection, and review of the manuscript.

Naser Alali: study design, data collection, and review of the manuscript.

Katy Dmowski: study design, data collection, and review of the manuscript.

Pepa Kaneva: study design, coordination of ethics approval, review of the manuscript.

Nicolò Pecorelli: study design, data interpretation, and review of the manuscript.

Liane S. Feldman: study design, supervision of recruitment and data collection, review of the manuscript.

Marylise Boutros: study design, supervision of recruitment and data collection, review of the manuscript.

Lawrence Lee: study design, supervision of recruitment and data collection, review of the manuscript.

Julio F. Fiore Jr (Primary supervisor): study design, project coordination, supervision of recruitment, data collection, statistical analysis, data interpretation, review of manuscript and thesis.

LIST OF TABLES & FIGURES

Figure 1. Postoperative recovery trajectory

Figure 2. Participant flowchart

Table 1. Definitions of measurement properties

Table 2. Participants' demographics, operative characteristics, and postoperative outcomes (n=282).

Table 3. PROMIS-29 physical and mental health summary scores, internal consistency, and floor and ceiling effects.

Table 4. Content validity of PROMIS-29.

Table 5. Construct validity of PROMIS-29 physical and mental health summary score

Table 6. Responsiveness of PROMIS-29 physical and mental health summary score.

LIST OF ABBREVIATIONS

ASA:	American Society of Anesthesiologists
ASER:	American Society for Enhanced Recovery
BMI:	Body Mass Index
CAT:	Computer Adaptive Testing
CCI:	Charlson Comorbidity Index
CD:	Clavien-Dindo
COSMIN:	CONsensus-based Standards for the selection of health status Measurement INstruments
ED:	Emergency Department
ERAS:	Enhanced Recovery After Surgery
ERPs:	Enhanced Recovery Pathways
FDA:	Food and Drug Administration
HRQoL:	Health-Related Quality of Life
HR-PRO:	Health-Related Patient-Reported Outcome
ICF:	International Classification of Functioning, Disability, and Health
IRT:	Item Response Theory
ISOQOL:	International SOciety for Quality of Life Research
JGH:	Jewish General Hospital
LOS:	Length of Stay
MGH:	Montreal General Hospital
MHS:	Mental Health Summary
MIS:	Minimally invasive surgery
MID:	Minimal Important Difference
NIH:	National Institutes of Health
PACU:	Post-Anaesthesia Care Unit
PHS:	Physical Health Summary
POQI:	PeriOperative Quality Initiative
POW:	PostOperative Week
PRO:	Patient-Reported Outcomes
PROMIS:	Patient-Reported Outcomes Measurement Information System
PROMs:	Patient-Reported Outcomes Measures

SEM: Standard error of measures

SD: Standard deviation

SSD: Same-day Discharge

WHO: World Health Organization

CHAPTER 1 -INTRODUCTION

1.1 Recovery After Colorectal Surgery

Surgery offers symptom relief or a potential cure for many malignant and benign colorectal conditions, including cancer, diverticulitis, and inflammatory bowel disease [1–6]. Colorectal surgery represents a considerable volume of operating room procedures, with approximately 300,000 colorectal resections conducted in North America every year [7, 8]. Despite its therapeutic benefits, colorectal surgery imposes substantial physiological stress on patients, leading to a cascade of metabolic and hormonal events triggered by surgical trauma [9]. This surgical stress response, characterized by systemic inflammation, immunosuppression, and insulin resistance, invariably leads to a rapid decline in postoperative health – manifested as pain, fatigue, reduced mobility, and impaired bowel function – that requires weeks to months for full recovery (**Figure 1**) [9, 10].

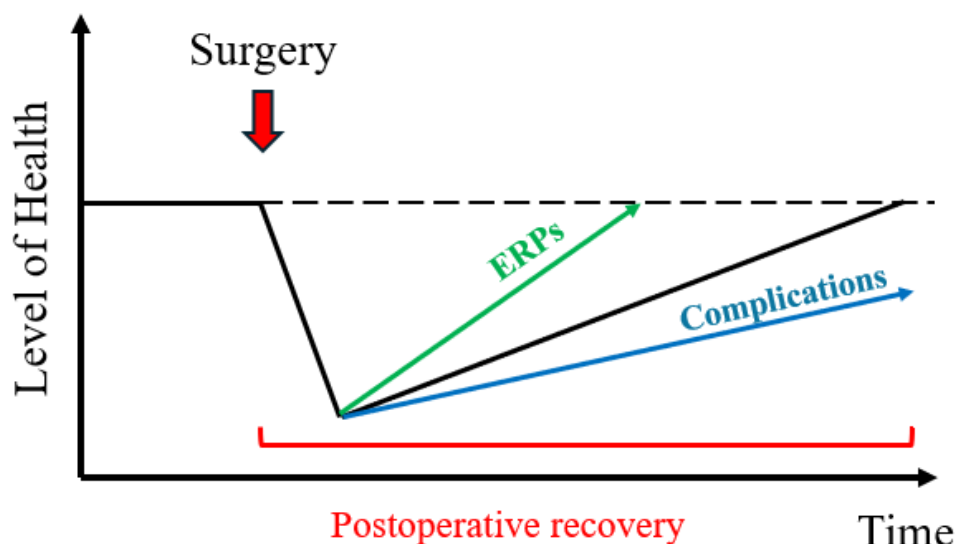


Figure 1. Postoperative recovery trajectory

The degree of surgical stress experienced by patients positively correlates with the extent of surgical trauma and with their likelihood of developing postoperative complications [11–13]. In fact, colorectal surgical procedures are associated with a significant morbidity burden, with approximately 30% of patients experiencing postoperative complications [14] such as postoperative ileus, infections, and anastomotic leaks [15, 16]. Given these complications, many patients require prolonged hospital stays (>5 days), post-discharge emergency department visits (~20%), and hospital readmissions (~10%) [17–19]. Importantly, complications also considerably prolong postoperative recovery and may prevent some patients from fully restoring their preoperative level of function. Prolonged or incomplete recovery not only increases direct medical costs [20], but also imposes substantial indirect burdens on patients, carers, and society, including lost work time, reduced wages, diminished productivity, and a lower quality of life [21].

1.2 Perioperative Care in Colorectal Surgery

Given the growing recognition of postoperative morbidity as a major factor affecting patients and healthcare systems, improving recovery after colorectal surgery has become a crucial target for quality-of-care improvement [22]. To address this issue, Enhanced Recovery Pathways (ERPs) have gained widespread acceptance as the new standard of care after colorectal surgery [23]. Introduced by Henrik Kehlet et al. in the 1990s [24], ERPs aim to minimize surgical stress response and enhance recovery through comprehensive, multidisciplinary care involving multiple preoperative, intraoperative, and postoperative interventions [25]. In colorectal surgery, ERAS® society guidelines recommend 24 interventions as part of patients' 'optimal' perioperative care, including patient education and engagement, minimally invasive surgery, multimodal analgesia, early oral nutrition, and mobilization [25, 26].

Over the past three decades, innovations in surgical technologies have also been proposed to improve postoperative recovery outcomes. Minimally invasive surgery (MIS) performed using laparoscopic, hand-assisted, and robotic approaches, has become the preferred approach for performing colorectal surgical procedures. Compared to open surgery, these techniques reduce blood loss, postoperative pain intensity, and facilitate earlier recovery of bowel function and mobilization [27]. Furthermore, MIS has shown benefits in reducing patients' length of hospital stay, postoperative complications, and healthcare costs, particularly when combined with ERPs [23, 28–30]. A meta-analysis of randomized controlled trials conducted by Zhuang et al [31], demonstrated that, within ERP settings, MIS significantly decreases morbidity by 22% and total hospital stay by an average of 1.92 days, when compared to open surgery.

1.3 Measurement of Postoperative Recovery

1.3.1 Traditional Measures

Most studies on strategies aimed to enhance recovery after surgery rely on traditional outcome measures such as length of stay (LOS) and postoperative complications to evaluate intervention effectiveness [32, 33]. These metrics are clinician- or observer-reported outcomes that are relevant for surgeons, institutions, and payers [34]. Surgeons consider LOS as an index of a “successful” surgery, reflecting an effective procedure without complications [35]. To institutions and payers, LOS is an index of resource utilization as prolonged hospital stay may increase healthcare costs and impact the availability of beds in surgical wards, negatively affecting surgery waitlists [36, 37]. Postoperative complications are a relevant outcome for all stakeholders in surgery. They are undesirable to patients, can impact surgeons' reputation, and lead to unplanned healthcare re-utilization (i.e., emergency department visits, and readmissions). Importantly, in countries such as the United States,

hospitals may face reduced reimbursements due to penalties associated with high complication-related readmission rates [38]. However, despite the undeniable importance of LOS and complications as indicators of surgical success to care providers, these traditional metrics have limitations when used as surrogate measures of recovery.

For instance, LOS can be influenced by external factors unrelated to the recovery process, such as the absence of post-discharge family support, insurance status, or lack of established discharge destination [39, 40]. Postoperative complications are inconstantly defined across studies, making it challenging to compare the results from studies focused on this outcome [41–43]. Importantly, these traditional measures fail to fully capture the complexity of the recovery process, which encompasses multiple dimensions of health and extends beyond hospital discharge [10, 12, 44]. Finally, these metrics overlook the perspective of patients – those who are actually recovering from the surgical procedure [45, 46]. For patients, the meaning of recovery is multifaceted and encompasses numerous dimensions, such as overcoming mental challenges, resuming normal habits and routines, reclaiming independence, alleviating symptoms, and re-finding joy in life [45]. Therefore, to effectively measure postoperative recovery, there is a need for metrics that are centered on the patient's perspective and consider all dimensions of recovery.

1.3.2 Patient-reported Outcomes Measures (PROMs)

Patient-reported outcomes measures (PROMs) have emerged as valuable tools for measuring postoperative recovery. PROMs are standardized questionnaires completed by patients themselves, which provide insights into their personal health status and recovery experience [34]. They can be unidimensional or multidimensional, measuring single or multiple constructs [47, 48]. They can also be generic, addressing general aspects of health applicable to the entire population, or condition-specific, focusing on aspects affected by a

particular disease or condition [47, 48]. By incorporating PROMs into perioperative care research, patients are engaged as the key stakeholders in their recovery process [49]. Furthermore, PROMs have the potential to identify patients with persisting symptoms or postoperative functional impairments, which are important in guiding quality improvement initiatives [47]. When integrated into routine clinical practice, PROMs can help surgeons gain a deeper understanding of what recovery means to their patients. These insights can potentially enhance personalized care and increase patient satisfaction [34].

The American Society for Enhanced Recovery (ASER) and Perioperative Quality Initiative (POQI) workgroup, an international, multidisciplinary non-profit organization that sets standards for optimizing perioperative care, established a consensus on best practices for integrating PROM in the context of postoperative recovery [50]. Their recommendations include collecting questionnaires at different time points, specifically preoperatively (at baseline), immediately after surgery, and after hospital discharge, to delineate patients' recovery trajectories. Additionally, questionnaires with a shorter recall period (less than 24h) should be used within the first week after surgery (including in-hospital stay), while those with a longer recall period should be preferred for evaluating long-term recovery beyond the first week [50].

Many studies have evaluated the impact of interventions to enhance recovery (i.e., ERPs and MIS) using PROMs [51, 52]. However, their findings were conflicting, suggesting that while these interventions improve LOS and complication rates, they may not necessarily improve recovery from the patients' perspective [51, 52]. This uncertainty raises questions about whether the lack of patient-reported improvement reflects a genuine failure of the interventions to enhance recovery or if it indicates that the PROMs used were inadequate for measuring recovery. To draw valid conclusions, it is essential that PROMs are supported by strong psychometric properties.

1.3.3 Psychometric Properties of PROMs

Measuring recovery through PROMs requires a careful selection of questionnaires that are capable of capturing aspects of perioperative health status that are meaningful to patients. This ensures that conclusions drawn from PROM data are robust and relevant. Consequently, establishing the psychometric properties of PROMs within the context of recovery is of paramount importance. Psychometric properties refer to the characteristics and qualities of PROMs, including assessments that determine their quality. According to the Consensus-based Standards for the Selection of Health Status Measurement Instruments (COSMIN) [53], nine psychometric properties are relevant to PROMs: content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, construct validity, and responsiveness. To the International Society for Quality-of-Life Research (ISOQOL), robust evidence regarding four of these properties is minimally required for the use of PROMs in studies on comparative-effectiveness: content validity, internal consistency, construct validity, and responsiveness (definitions in **Table 1**) [54].

Measurement property	Definitions
<i>Content Validity</i>	The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured.
<i>Internal Consistency</i>	The degree of interrelatedness among items.
<i>Construct Validity</i>	The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the HR-PRO instrument validly measures the construct to be measured.
<i>Responsiveness</i>	The ability of an HR-PRO instrument to detect change over time in the construct to be measured.
<i>HR-PRO</i> , health-related patient-reported outcome.	

Table 1. Definitions of measurement properties [55].

According to COSMIN, content validity is the single most important psychometric property of a PROM [56]. To evaluate the content validity in the context of recovery, the first essential step is to comprehensively understand recovery from the patients' perspective. In fact, the Food and Drug Administration (FDA) guidelines for PROMs emphasize the importance of identifying themes that best describe the concept of interest (i.e., recovery) through qualitative studies and incorporating them into a conceptual framework [57]. In line with these principles, Fiore et al. [46] conducted a qualitative study involving 30 patients from four different countries, identifying 39 recovery-related themes and establishing the first conceptual framework of recovery after abdominal surgery [46].

It is important to note that content validity can impact other psychometric properties [56]. For instance, asking questions that do not fit a context-specific conceptual framework (e.g., enquiring patients about knee pain intensity after abdominal surgery) may reduce internal consistency-reliability. Not having sufficient items to cover all relevant aspects of a

conceptual framework (e.g., not asking about the re-establishment of gastrointestinal function after abdominal surgery) can negatively impact other fundamental psychometric properties such as construct validity and responsiveness [56]. This hinders the PROM's ability to differentiate between patients with different recovery patterns or detect changes over time.

Although PROMs should have robust psychometric properties when used to measure postoperative recovery, a comprehensive systematic review conducted by Fiore et al. identified important research gaps in this field. The review addressed the psychometric properties of 22 PROMs used to measure recovery after abdominal surgery [58] and supported that none of these PROMs met the minimum methodological standard set by ISOQOL [54]. Specifically, several PROMs showed limited content validity and the quality of studies on other psychometric properties was generally limited [58].

As an ideal recovery-related PROM is yet to be determined, while new research emerges, the ASER/POQI workgroup has recommended some PROMs for use in both research and clinical practice [50]. These recommendations, based on expert consensus, include questionnaires from the National Institutes of Health (NIH) Patient-Reported Outcome Measurement Information System (PROMIS), such as the PROMIS-29 [50].

1.3.4 Patient-Reported Outcome Measurement Information System (PROMIS)

PROMIS was established in 2004 as one of the initiatives of the NIH Roadmap for Medical Research, which aimed to develop PROMs according to optimal standards for use in research and routine clinical practice [59, 60]. This system comprises an extensive array of generic PROMs, targeting domain-specific or multidimensional aspects of health [61, 62]. PROMIS PROMs were created using Item Response Theory (IRT), a technique that uses diagnostic information (e.g., error estimates and fit statistics) to evaluate whether adding the

scores from a collection of items is justified [63–65]. By providing detailed insights into the properties of individual PROM items, IRT enhances the precision of measuring patients' health state [63–65]. This precision allows for the development of adaptive questionnaires that adjust to each patient's responses, thereby reducing the burden on patients while maintaining high measurement accuracy [66, 67] PROMIS measures are calibrated against norms from the US population [68] and their scores are standardized to a common metric (T-score with mean=50 and SD=10), enhancing comparability of health states across different patient populations [69].

The PROMIS profiles consist of a series of tools designed to evaluate key health domains through fixed short forms or computer adaptive tests. Among the available instruments, PROMIS-29, PROMIS-43, and PROMIS-57 differ primarily in the number of items per domain, with 4, 6, and 8 items, respectively [70]. While longer profiles, such as PROMIS-43 and PROMIS-57, offer greater reliability and precision, they may risk overburdening patients, making them less practical for routine clinical use. Conversely, PROMIS-29 strikes a balance by reducing patient burden while maintaining coverage across essential domains[70].

The PROMIS-29 adult profile is a generic multidimensional questionnaire that addresses seven health domains (i.e., physical function, fatigue, pain interference, depression, anxiety, ability to participate in social roles and activities, and sleep disturbance). Each domain consists of four items selected from the PROMIS item banks and calibrated using IRT [71–73]. The questionnaire also includes a single 0–10 numeric rating to assess pain intensity [70, 74]. Recent research by Hays et al. [75] supports that PROMIS-29 can be further summarized in one physical and one mental health summary score by combining specific domain scores.

1.4 Research gap and study aim

By assessing general aspects of physical and mental health, PROMIS-29 may represent a suitable tool for measuring recovery after surgery. However, despite having its use recommended by consensus[50], evidence regarding the psychometric properties of PROMIS-29 in the context of perioperative care remains scarce. As patients undergoing colorectal surgery have a considerable postoperative morbidity burden, they are an ideal population for studies assessing the psychometric properties of PROMs aimed at assessing recovery. Therefore, this thesis research aimed to contribute evidence regarding the content validity, internal consistency reliability, construct validity, and responsiveness of PROMIS-29 in the context of recovery after colorectal surgery.

CHAPTER 2 -MANUSCRIPT

Submitted at Annals of Surgery (November 2024)

Psychometric assessment of PROMIS-29 as a measure of recovery after colorectal surgery

Running head: PROMIS-29 in colorectal surgery

Francesca Fermi MD^{1,2,3}, Samin Shirzadi MD MPH^{2,3}, Ghadeer Olleik MSc^{1,2,3}, Makena Pook BHSc^{1,2,3}, Maxime Lapointe Gagner MSc^{1,2,3}, Sarah Al Ben Ali MD^{1,2}, Philip Nguyen-Powanda MSc^{1,2,3}, Elahe Khorasani PhD^{1,2,3}, Christos Mousoulis MD PhD^{1,2,3}, Tahereh Najafi Ghezeljeh PhD^{2,3}, Naser Alali MD^{1,2}, Katy Dmowski BSc^{1,2,3}, Pepa Kaneva MSc^{2,3}, Nicolò Pecorelli MD MSc⁴, Marylise Boutros MD^{1,5}, Lawrence Lee MD PhD^{1,3}, Liane S. Feldman MD^{1,3}, Julio F. Fiore Jr PhD^{1,2,3}

¹ *Department of Surgery, McGill University, Montreal, QC, Canada.*

² *Perioperative Care and Outcomes Research (PCOR) Lab, McGill University, Montreal, QC, Canada.*

³ *Centre for Outcomes Research and Evaluation (CORE), Research Institute of the McGill University Health Centre, Montreal, QC, Canada.*

⁴ *Vita-Salute San Raffaele University, Milan, Italy*

⁵ *Division of Colon and Rectal Surgery, Jewish General Hospital, Montreal, QC, Canada.*

Keywords Colorectal surgery, Patient-reported outcomes, Measurement, Recovery.

Funding information: Dr. Fiore and Dr. Lee receive salary support from the Fonds de Recherche du Québec-Santé (FRQS).

2.1 Introduction

Despite important advances in perioperative care (i.e., minimally invasive surgery[76] and enhanced recovery pathways [ERPs][25]), colorectal surgery remains associated with a substantial morbidity burden [14] Postoperative complications affect 30-50% of patients [77] and often result in functional impairment, requiring weeks to months for full recovery[12]. Therefore, there is an urgent need to optimize strategies to measure and improve recovery after colorectal surgery [78, 79].

Research focused on interventions aimed at enhancing postoperative recovery often relies on traditional measures such as length of hospital stay (LOS) or complication rates to assess treatment outcomes [32]. However, these measures fall short of capturing the complexity of the recovery process or patients' perspective [10, 32]. Recognizing this limitation, recent literature advocates for the measurement of postoperative recovery using patient-reported outcomes (PROs)[80, 81] – reports of health provided directly by the patient without interpretation by others [34]. Nevertheless, PRO measures (PROMs) currently used in the context of recovery after colorectal surgery were generally not developed according to robust scientific standards and lacked adequate psychometric properties (i.e., evidence of quality and precision)[58].

The questionnaire PROMIS-29 [70], derived from the National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS)[82], was developed according to rigorous qualitative and quantitative methods, using an item response theory (IRT) framework to optimize item selection and create unidimensional linear scores [83]. As PROMIS-29 measures general aspects of physical and mental health that may be relevant to surgical recovery, consensus recommendation endorses its use in the context of perioperative care and research[50]. However, despite this endorsement, evidence regarding the psychometric properties of PROMIS-29 in the context of colorectal surgery remains

scarce[58]. To address this knowledge gap, this study aimed to assess the content validity, internal consistency reliability, construct validity, and responsiveness of PROMIS-29 in the context of recovery after colorectal surgery.

2.2 Material and Methods

Study design and setting

This psychometric study was designed according to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)[84]. We analysed secondary data from a cohort study focused on opioid consumption and health status following colorectal surgery[85]. Ethics approval was obtained from the participating institutions (MUHC REB ref.: MP-37-2020-6273, CIUSSS West-Central REB ref: MEO-37-2020-2156), and all participants provided informed consent.

Eligible study participants were adult patients (≥ 18 years) undergoing elective colorectal resection at two university-affiliated hospitals in Montreal, Canada, between February 2021 and September 2022. Consistent with the primary study's selection criteria, we excluded patients (1) undergoing major concomitant non-colorectal procedures (e.g., liver resection); (2) with conditions hindering PRO assessment (e.g., cognitive impairment, inability to understand English or French); (3) with anticipated difficulties in post-discharge follow-up (e.g., limited access to a telephone or computer). To reduce sample heterogeneity, in this secondary analysis, our cohort was further refined by excluding patients who did not receive a bowel resection (e.g., ostomy reversals). At both participating institutions, patients received perioperative care following enhanced recovery after surgery (ERAS) principles, with a targeted length of stay (LOS) ≤ 3 days[25]. In one of the institutions, selected patients were eligible for same-day discharge[86].

Measures and procedures

We reviewed electronic medical records to obtain data concerning patient characteristics (i.e., age, sex, surgical diagnosis, Body Mass Index [BMI; kg/m²], comorbidities, American Society of Anaesthesiologists [ASA] score), procedure characteristics (i.e., type of surgery, surgical approach, procedure length, stoma creation) and postoperative outcomes (i.e., LOS, 30-day complications [classified according to Clavien-Dindo [43]], 30-day emergency department [ED] visits, and 30-day hospital readmissions).

The PROMIS-29 [70] was administered preoperatively and at postoperative weeks (POW) 1, 2, 3, and 4. This 29-item questionnaire targets seven health domains, including anxiety, depression, fatigue, pain interference, physical function, sleep disturbance, and participation in social roles and activities. Responses are obtained on a 5-point Likert scale, with an additional 0–10 numerical scale to quantify pain intensity[70]. Each domain consists of four items targeting a seven-day recall period, except the domains ‘physical function’ and ‘participation’, where the recall period is not specified. PROMIS-29 domain scores can be combined into a Physical Health Summary Score (PHS) and Mental Health Summary Score (MHS) using algorithms developed by Hays et al.[75] All scores are expressed as t-scores, with 50 as the mean score $\pm$ 10 standard deviations, referencing a US population [68]. T-scores can be calculated for a domain if at least one item of that domain is completed, with *data* deemed ‘missing’ if no items are completed [87]. Higher scores indicate more of the concept being measured. Our primary psychometric analysis focused on PROMIS-29 PHS and MHS. Secondary analyses examined individual PROMIS-29 health domains.

Content validity assessment

Content validity refers to the degree to which the content of a PROM questionnaire provides an adequate reflection of the construct to be measured [55]. In the present study, the construct of interest is ‘postoperative recovery’ in the context of colorectal surgery. To assess content validity, PROMIS-29 items were evaluated against the conceptual framework of

recovery after abdominal surgery proposed by Alam et al. [46] This framework has been informed by qualitative interviews with patients undergoing abdominal surgery (including colorectal) in four different countries (Canada, Brazil, Italy, and Japan). Both PROMIS-29 items and framework domains[46] have been previously mapped to the World Health Organization's International Classification of Functioning, Disability, and Health (ICF)[88], which facilitates content validity assessment by providing a common language for describing health-related states[89]. To appraise content validity, we evaluated PROMIS-29's relevance and comprehensiveness in relation to the conceptual framework of recovery after abdominal surgery[56]. To evaluate relevance (are all PROMIS-29 items relevant to the process of postoperative recovery?), we assessed the extent to which PROMIS-29 items align with the recovery framework[56]. To evaluate comprehensiveness (are all framework domains addresses by PROMIS-29?), we assessed the extent to which the framework domains are covered by PROMIS-29 items[56].

Content validity was further appraised by calculating floor and ceiling effects [55, 56]. These occur when a substantial number of respondents (>15%) score at the lower or upper extremes of a questionnaire, respectively[90]. With floor or ceiling effects, the questionnaire may not be sensitive enough to detect differences at both ends of the scale. The maximum and minimum t-score values for each specific PROMIS-29 domain were obtained from the questionnaire's scoring manual (version 2.1)[70]. Based on these values, we calculated floor and ceiling effects for summary scores (PHS and MHS) using the formula by Hays et al. [75]

Internal consistency reliability

Internal consistency is a crucial component of reliability assessment, examining the degree of interrelatedness among the items in a PROM[55]. In our evaluation of internal consistency, we employed Cronbach's alpha to assess the extent to which PROMIS-29 items measure a single underlying trait within each summary and domain scale[55].

Construct validity assessment

Construct validity refers to the degree to which the scores of a PROM are consistent with hypotheses based on the assumption that the questionnaire validly measures the construct of interest (i.e., postoperative recovery)[55]. The construct validity of PROMIS-29 scores was assessed by testing *a priori* hypotheses based on the assumption that, across the four postoperative time points, scores would indicate ‘better’ recovery in patients:

- 1) with shorter length of stay ($\text{LOS} \leq 4$ days) vs. longer ($\text{LOS} > 4$ days)[91, 92].
- 2) without 30-day postoperative complications vs. with 30-day complications[91–93].
- 3) undergoing minimally invasive surgery vs. open (or converted)[92, 94].
- 4) without a new stoma vs. with a new stoma [94, 95].

Construct validity was deemed 'sufficient' when at least 75% of the construct validity hypotheses were supported by the data[53]. In a *post hoc* analysis, we assessed PROMIS-29 scores among patients without 30-day postoperative severe complications vs. with 30-day postoperative severe complications (Clavien-Dindo grade III or above)[91–93].

Responsiveness assessment

Responsiveness is the ability of a PROM to detect changes over time in the construct being measured[55]. This was assessed by testing *a priori* hypotheses based on the assumption that PROMIS-29 scores follow the expected trajectory of recovery after surgery. This trajectory is characterized by an initial decline in patients’ health status followed by a gradual return toward baseline (preoperative) health [10]. Therefore, we hypothesized that PROMIS-29 scores would indicate that patients have a ‘better’ health status:

- 1) preoperatively vs. at POW1 (decline expected).
- 2) at POW2 vs. POW1 (improvement expected).
- 3) at POW3 vs. POW2 (improvement expected).

4) at POW4 vs. POW3 (improvement expected).

Responsiveness was deemed 'sufficient' when at least 75% of the responsiveness hypotheses were supported by the data[96].

Sample size

This study used secondary data from 282 patients enrolled in a cohort study. According to COSMIN standards, studies including ≥ 100 participants are considered to have an adequate sample size for the assessment of internal consistency, construct validity, and responsiveness [97]. The content analysis conducted in our study is descriptive and not sensitive to sample size.

Statistical analysis

All statistical analyses were conducted using Stata Version 18 software (StataCorp, College Station, TX, USA). Categorical variables were presented as frequencies and percentages, while continuous variables were expressed as means and standard deviations (SD) or medians with interquartile ranges, as appropriate. Data concerning content validity were analyzed using descriptive statistics. For internal consistency assessment, Cronbach's alpha was calculated for PROMIS-29 summary and domain scores at each time point[98, 99]. Linear regression was employed to test construct validity and responsiveness hypotheses with PROMIS-29 scores serving as the dependent variable. In the construct validity analysis, regression coefficients represent the between-group mean difference in PROMIS-29 scores, after adjustment for baseline scores. To test responsiveness, we ran different linear regression models for each specific hypotheses (baseline vs. POW1, POW2 vs. POW1, POW3 vs. POW2, POW4 vs. POW3) and regression coefficients represent the mean difference in PROMIS-29 scores between the time points of interest. Missing data were addressed using

multiple imputations with predictive mean matching [100]. Estimates from 50 imputed datasets were combined using Rubin's rules [101].

According to COSMIN recommendations, it is advised that psychometric hypothesis testing be based on the expected direction and magnitude of differences/relationships, rather than p-values (which can be influenced by sample size and multiple comparisons)[102]. For internal consistency, we hypothesized that Cronbach's alpha would be acceptable (≥ 0.7 [90]). For construct validity and responsiveness, we hypothesized that differences between groups or time points would exceed a minimal important difference (MID) of 3.0 t-score points. This MID value is supported by systematic review findings [103] and PROMIS leadership consensus [104]. To ensure the robustness of our results, we conducted sensitivity analyses considering MIDs calculated using distribution-based methods: (1) standard error of measures (SEM) determined as the SD multiplied by square root of 1-Cronbach's and (2) based on half of a SD of each baseline score [105].

2.3 Results

A total of 755 patients were screened during the study period and 282 were included in the psychometric analysis. The study flow diagram and reasons for non-participation are described in **Figure 2**. Most patients (73%) completed all the questionnaires during the follow-up period. Rates of missing data at different time points (addressed using multiple imputation) ranged from 2-14% (**Supplemental Digital Content 1**).

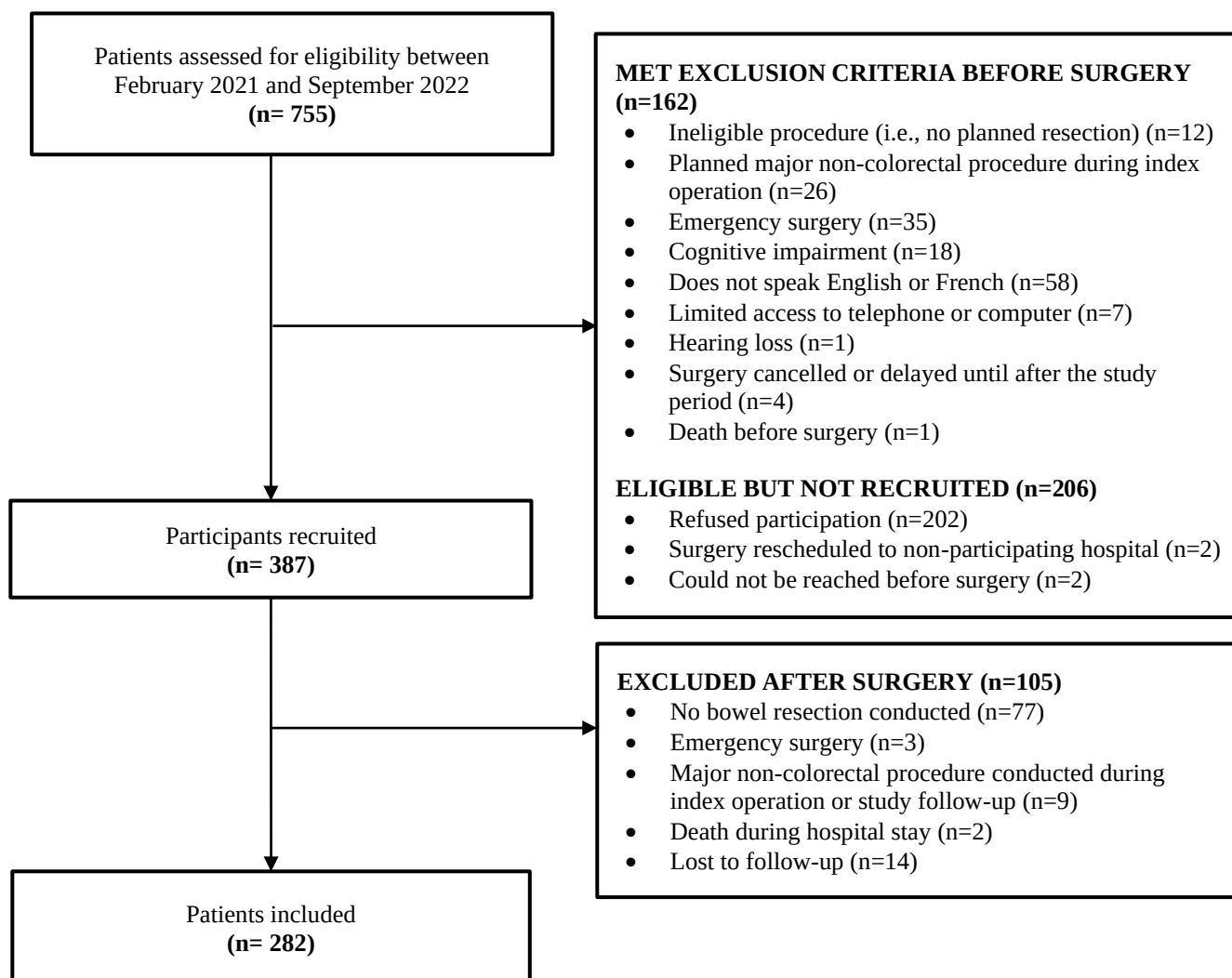


Figure 2. Participant flowchart

The median age of participants was 59 ± 16 years, 52% were male, 64% had ASA score ≤ 2 , 61% had surgery for malignancy, and 19% received a new stoma (**Table 2**). Most participants underwent minimally invasive surgery (78%) and colon resection (62%). The median LOS was 3 days [IQR 1-6] and 16% of participants were discharged on the same day of the procedure. Within 30 days, 34% of participants developed a postoperative complication (8% major complications), 20% required an ED visit, and 9% were readmitted to the hospital. Rates and definitions of specific complications are reported in **Supplemental Digital Content 2**.

Table 2. Participants' demographics, operative characteristics, and postoperative outcomes (*n*=282).

Age, years	59 ± 16
Sex (Male)	148 (52)
BMI, Kg/m²	26 [23-29]
ASA score	
I-II	179 (64)
III-IV	101 (36)
Charlson Comorbidity Score	2 [1-3]
Diagnosis	
Malignancy	173 (61)
Diverticulitis	32 (11)
Inflammatory bowel disease	48 (17)
Others	29 (10)
Surgical approach ^a	
Minimally invasive	220 (78)
Open	62 (22)
Type of surgery	
Colon ^b	175 (62)
Rectal ^c	107 (38)
Formation of new stoma	55 (19)
Time of Surgery, min	215 [163-308]
30-day complication	97 (34)
Minor complication (CD I-II)	260 (92)
Major complication (CD III-V)	22 (8)
30-day emergency department visits	55 (20)
30-day hospital readmissions	25 (9)
Length of stay, days	3 [1-6]
Same day discharge	46 (16)

Data are expressed as n (%), mean ± SD, median [interquartile range].

BMI, Body Mass Index; *ASA*, American Society of Anaesthesiologists; *CD*, Clavien-Dindo.

^a Surgical approach was categorized into open (including converted [*n*=22]) and minimally invasive surgery (including laparoscopic [*n*=217] and hand-assisted [*n*=3]).

^b Procedures included are small bowel resection (*n*=14), right hemicolectomy (*n*=73), ileocolic resection (*n*=20), transverse colectomy (*n*=2), left hemicolectomy (*n*=16), sigmoid resection (*n*=27), subtotal/total colectomy (*n*=9), Hartmann's procedure (*n*=4), reversal Hartmann's procedure (*n*=10).

^c Procedures included are anterior rectal resection (*n*=83), abdominoperineal resection (*n*=12), proctocolectomy (*n*=7), and proctocolectomy with ileal-pouch anal anastomosis (*n*=5).

Data concerning PHS and MHS at different time points are reported in **Table 3**. Data concerning domain-specific PROMIS-29 scores are reported in **Supplemental Digital Content 3**.

Table 3. PROMIS-29 physical and mental health summary scores, internal consistency, and floor and ceiling effects.

	Sample (n)	Mean $\pm$ SD	Median [IQR]	Cronbach Alpha	Floor Effect (%)	Ceiling Effect (%)
<i>Physical health summary score (range 20.4-58.9)</i>						
Preoperatively	274	49 $\pm$ 10	55 [42-58]	0.89	0	0
POW1	252	40 $\pm$ 8	39 [35-44]	0.88	0	0
POW2	250	44 $\pm$ 8	43 [39-49]	0.88	0	0
POW3	243	46 $\pm$ 9	46 [40-56]	0.91	0	0
POW4	256	49 $\pm$ 9	49 [43-58]	0.90	0	0
<i>Mental health summary score (range 19.5-66.1)</i>						
Preoperatively	274	50 $\pm$ 9	50 [44-56]	0.89	0	0
POW1	252	47 $\pm$ 8	47 [42-52]	0.88	0	0
POW2	250	50 $\pm$ 8	50 [45-56]	0.88	0	0
POW3	243	51 $\pm$ 9	52 [46-58]	0.91	0	0
POW4	256	53 $\pm$ 9	54 [48-60]	0.90	0	0

Higher PHS scores indicate better physical health; Higher MHS scores indicate better mental health.

The floor effect represent the frequency of patients achieving the lowest PROMIS-29 t-score.

The ceiling effect represent the frequency of patients achieving the highest PROMIS-29 t-score.

The minimally important difference (MID) was calculated as half of the standard deviation of each baseline t-score.

POW, postoperative week

Psychometric analysis

Content Validity

The content analysis of PROMIS-29 items in relation to the conceptual framework of recovery after abdominal surgery is summarized in **Table 4**. Overall, most of the ICF domains covered by PROMIS-29 (14/15, 93%) are pertinent to the conceptual framework of recovery after abdominal surgery, revealing high relevance. This indicates that the content of PROMIS-29 is generally relevant to the process of postoperative recovery, particularly concerning the recovery of overall physical and mental health. However, many ICF domains included in the conceptual framework (15/38, 39%) are not covered by PROMIS-29, indicating low comprehensiveness. This suggests that PROMIS-29 does not cover many issues that are relevant to the process of postoperative recovery, including recovery of visceral functions (i.e., digestive, defecation, and urinary functions), skin repair, ability to change body positions, and ability to self-care.

Table 4. Content validity of PROMIS-29.

ICF domains covered by PROMIS-29 or the Conceptual Framework of Recovery			
First-level domains	Second-level categories	Covered by the Conceptual Framework?	Covered by PROMIS-29?
Body functions (b)			<i>Items (n)</i>
Mental functions (b1)	Energy and drive functions (130)	✓	✓ 4
	Sleep functions (134)	✓	✓ 4
	Attention functions (140)	✓	✓ 1
	Emotional functions (152)	✓	✓ 8
Sensory functions and pain (b2)	Sensations associated with hearing and vestibular functions (240)	✓	X 0
	Sensation of pain (280)	✓	✓ 5
Functions of the cardiovascular, hematological, immunological, and respiratory systems (b4)	Respiratory functions (440)	✓	X 0
	Exercise tolerance functions (455)	✓	✓ 2
Functions related to the digestive, metabolic, and endocrine systems. (b5)	Ingestion functions (510)	✓	X 0
	Digestive functions (515)	✓	X 0
	Defecation functions (525)	✓	X 0
	Sensation associated with the digestive system (535)	✓	X 0
Genitourinary and reproductive functions (b6)	Urination functions (620)	✓	X 0
Neuromusculoskeletal and movement-related functions (b7)	Muscle power functions (730)	✓	X 0
Functions of the skin and related structures (b8)	Repair functions of the skin (820)	✓	X 0
	Sensation related to the skin (840)	✓	X 0
Activities and participation			
General tasks and demands (d2)	Undertaking multiple tasks (220)	✓	X 0
	Carrying out daily routine (230)	✓	✓ 1
	Handling stress and other psychological demands (240)	✓	X 0
Mobility (d4)	Changing basic body position (410)	✓	X 0
	Maintaining body position (415)	✓	X 0
	Lifting and carrying objects (430)	✓	X 0
	Hand and arm use (445)	✓	X 0
	Walking (450)	✓	✓ 1
	Moving around (455)	✓	✓ 1
	Using transportation (470)	✓	X 0
	Driving (475)	✓	X 0
Self-care (d5)	Washing oneself (510)	✓	X 0
	Caring for body parts (520)	✓	X 0
	Dressing (540)	✓	X 0
	Looking after one's health (570)	✓	X 0
Domestic life (d6)	Acquisition of goods and services (620)	✓	✓ 1
	Preparing meals (630)	✓	X 0
	Doing housework (640)	✓	✓ 2
	Assisting others (660)	✓	✓ 1
Interpersonal interactions and relationships (d7)	Basic interpersonal interactions (710)	X	✓ 1
	Intimate relationships (770)	✓	X 0

Major life areas (d8)	Remunerative employment (850)	✓	✓	1
Community, social, and civic life (d9)	Recreation and leisure (920)	✓	✓	1

The PROMIS-29 summary scores, PHS and MHS, did not present floor or ceiling effects as no patients achieved maximum or minimal scores at any time point (**Table 3**). In contrast, floor and ceiling effects (>15%) were common for PROMIS-29 domain scores, with the highest percentages of maximal and minimal scores observed preoperatively and at 4 weeks after surgery (**Supplemental Digital Content 3**). The domain ‘physical function’ exhibited the highest ceiling effect, with 53% of participants reporting the maximum level of physical function preoperatively and 40% at 4 weeks. The domain ‘depression’ showed the highest floor effect, with over 40% of participants reporting the lowest possible level of depression at all time points.

Internal Consistency Reliability

The internal consistency reliability across all PROMIS-29 summary and domain scores was deemed ‘good’ to ‘excellent’, with Cronbach’s alpha ranging from 0.81 to 0.98 (**Table 3; Supplemental Digital Content 3**).

Construct validity

The construct validity of PROMIS-29 summary scores in the context of recovery after colorectal surgery was limited. None of the *a priori* hypotheses concerning PROMIS-29 PHS scores were supported by the data on POW1, while only 1 out of 4 hypotheses (25%) were supported on POW 2, 3, and 4 (**Table 5**). For PROMIS-29 MHS scores, none of the *a priori* hypotheses were supported on POW1, 2, and 4, and only 1 hypothesis was supported on POW 3 (25%). Construct validity was not considerably improved when using an alternative hypothesis concerning postoperative complications (severe complications vs. no severe complications) (**Table 5**). Limited construct validity was also observed in sensitivity analyses

targeting a MID = SEM (zero PHS hypotheses supported on POW1 and 1 hypotheses supported on POW2, 3, 4; zero MHS hypotheses supported on POW1 and 2, 2 hypotheses supported on POW3, and 1 hypothesis supported on POW4). When considering a MID = 1/2 SD, no PHS and MHS construct validity hypotheses were supported (**Supplemental Digital Content 4 and 5**).

Table 5. Construct validity of PROMIS-29 physical and mental health summary scores.

	POW1	POW2	POW3	POW4
<i>Physical health summary score</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.1 (0.9 to 3.1)	2.6 (0.7 to 4.6)	3.4 (1.6 to 5.5)	3.3 (1.4 to 5.3)
2. No 30- day complications vs. 30-day complications	2.3 (0.4 to 4.3)	2.4 (0.5 to 4.3)	2.7 (0.6 to 4.7)	2.9 (1.0 to 4.9)
3. Minimally invasive vs. open	1.8 (-0.4 to 4.1)	3.2 (1.1 to 5.4)	2.9 (0.6 to 5.2)	2.3 (0.1 to 4.5)
4. No new stoma vs. new stoma	1.4 (-0.9 to 3.7)	1.8 (0.4 to 4.0)	2.2 (0.2 to 4.6)	1.6 (0.6 to 3.9)
- No 30-day severe complications vs. 30-day severe complications	0.9 (-2.8 to 4.6)	1.7 (-1.9 to 5.5)	2.9 (1.0 to 6.8)	3.3 (0.3 to 6.9)
<i>Mental health summary score</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	0.2 (-1.8 to 2.1)	2.3 (0.4 to 4.1)	2.9 (1.0 to 4.9)	2.2 (0.3 to 4.2)
1. No 30- day complications vs. 30-day complications	1.7 (-0.3 to 3.6)	2.4 (0.6 to 4.2)	2.6 (0.6 to 4.6)	2.1 (0.2 to 4.0)
2. Minimally invasive vs. open	0.3 (-2.6 to 1.9)	1.2 (-0.9 to 3.2)	1.3 (-0.9 to 3.5)	1.1 (-1.1 to 3.3)
3. No new stoma vs. new stoma	1.7 (-0.6 to 4.0)	2.5 (0.4 to 4.6)	3.3 (1.0 to 5.5)	2.1 (0.2 to 4.3)
- No 30-day severe complications vs. 30-day severe complications	1.4 (-2.4 to 5.1)	2.3 (-1.0 to 5.7)	3.6 (-0.2 to 7.4)	3.0 (-0.6 to 6.6)

Data are the mean difference (95% CI) in the t-score.

Mean differences were considered relevant if above the minimally important difference (MID) of 3 points.

Hypothesis supported by the data are bolded.

LOS, length of stay; POW, postoperative week. Severe complications are defined as Clavien-Dindo ≥ 3.

The construct validity of PROMIS-29 specific domains was also limited, with only 0% to 50% of *a priori* hypotheses supported (**Supplemental Digital Content 6**). The performance of some domains (i.e., “fatigue”, “physical function”, and “ability to participate in social roles and activities”) was improved when analyses considered only severe complications and targeted a MID = SEM; at certain time points, 75% of the hypotheses were

confirmed (**Supplemental Digital Content 7**). No construct validity hypotheses were supported when targeting a MID = 1/2 SD (**Supplemental Digital Content 8**).

Responsiveness

In our primary responsiveness analysis, both PHS and MHS scores were able to delineate the expected recovery trajectory when measured from baseline to POW2, but differences observed on POW3 and POW4 were no longer meaningful (**Table 6**). The same results were observed in sensitivity analyses considering a MID = SEM and a MID = 1/2 SD (**Supplemental Digital Content 4 and 5**).

Table 6. Responsiveness of PROMIS-29 physical and mental health summary scores.

	Mean difference (95% CI)
<i>Physical health summary score</i>	
1. Baseline vs. POW1	9.9 (8.5 to 11.4)
2. POW2 vs. POW1	3.9 (2.5 to 5.2)
3. POW3 vs. POW2	2.7 (1.3 to 4.2)
4. POW4 vs. POW3	2.4 (0.9 to 3.9)
<i>Mental health summary score</i>	
1. Baseline vs. POW1	3.1 (1.6 to 4.5)
2. POW2 vs. POW1	3.1 (1.8 to 4.5)
3. POW3 vs. POW2	1.6 (0.2 to 3.0)
4. POW4 vs. POW3	1.8 (0.3 to 3.2)

Data are the mean difference (95% CI) in the t-score.

Hypotheses supported by the data are bolded.

Mean differences were considered relevant if above the minimally important difference (MID) of 3 points.

POW, postoperative week.

The responsiveness of PROMIS-29 specific domains was also limited, with 0% to 50% of the hypotheses supported by the data (**Supplemental Digital Content 9**). The domain “pain interference” was sensitive to alternative MIDs, with all the responsiveness hypotheses confirmed (100%) when considering a MID = SEM (**Supplemental Digital Content 10 and 11**).

2.4 Discussion

Contemporary literature underscores the importance of using PROMs to assess postoperative recovery from the patient's perspective[10, 58, 80]. The popularity of the PROMIS measurement system, proposed by the NIH, prompted guidelines to suggest the PROMIS-29 questionnaire as a potential tool for measuring patient-reported recovery after colorectal surgery[50]. However, the present study supports that PROMIS-29 has limited content validity, construct validity, and responsiveness in this context. While PROMIS-29 showed strong internal consistency, it only captured 39% of the themes that patients consider relevant to the process of recovery after abdominal surgery. Additionally, *a priori* hypotheses tested for construct validity had very limited support, underscoring the inability of PROMIS-29 to differentiate between patients with different recovery trajectories. Finally, PROMIS-29 was responsive to the expected recovery trajectory up to 2 weeks after surgery, but not thereafter.

This study's major strength lies in its adherence to COSMIN quality standards for the appraisal of PROMs[84]. To prevent reporting bias, our analyses focused on psychometric hypotheses posed *a priori*. All hypotheses were based on effect sizes (i.e., expected magnitude of differences and correlations) rather than statistical significance, thereby reducing the impact of sample size and type II error on our results. Furthermore, we analyzed a relatively large participant sample (n=282), in line with COSMIN sample size recommendations for studies on psychometric properties. Given these design considerations, we believe this study contributes important new knowledge to inform guidelines and future research on strategies to measure recovery after colorectal surgery.

Our results are in line with evidence supporting that generic patient-reported measures of health status, such as PROMIS-29, have limited comprehensiveness and sensitivity when used to measure recovery after colorectal surgery[58]. As with other generic PROMs such as the Short-form 36 (SF-36) and WHO Disability Assessment Schedule 2.0 (WHODAS

2.0)[58], PROMIS-29 does not cover many issues that are relevant to the process of postoperative recovery after abdominal surgery. These include the recovery of gastrointestinal, urinary functions, skin repair, the ability to change body positions, and the ability to self-care [46]. In addition, PROMIS 29 has a relatively long recall period (i.e. the period of time that patients are asked to consider when responding to PRO items) of 7 days. Given the highly dynamic nature of recovery in the first weeks following surgery, PROMs with extended recall periods may be unable to capture rapid changes in a patient's health status. These shortcomings may compromise the psychometric properties of PROMs, explaining the inability of PROMIS-29 to differentiate between patients expected to have different recovery patterns, such as those with and without postoperative complications, and its limited responsiveness to the expected recovery trajectory. Using PROMs not supported by strong psychometric properties can mislead clinicians and researchers, potentially resulting in research waste and misguided clinical decisions[106].

Our findings differ from those observed in a recent study by Pecorelli et al[91], which supported the construct validity and responsiveness of PROMIS-29 when measuring recovery up to 90 days after pancreatic surgery[91]. This discrepancy can potentially be attributed to the higher extent of physiological stress experienced by patients following pancreatic resections compared to colorectal resections. Surgical stress response correlates with surgical complexity and the extent of tissue damage[12], significantly impacting the risk of postoperative complications and the length of recovery[13]. Most patients in the present study underwent colonic resection, with a median LOS of 3 days (16% of patients having same-day discharge), and 8% experienced severe complications within 30 days. In Pecorelli's cohort [91], most patients underwent pancreaticoduodenectomy (53%), with a median LOS of 9 days, and 26% experienced 30-day severe complications. Given the higher morbidity associated with pancreatic procedures, patients' compromised postoperative health state may

further impact the PROMIS-29 domains, thereby influencing the questionnaire's psychometric properties. As with any abdominal surgery, many issues experienced by patients undergoing pancreatic resection are not addressed by PROMIS-29. However, Pecorelli et al[91] support that some recovery-related domains, such as overall physical and mental function, may be adequately captured by PROMIS-29 after complex pancreatic procedures. Our results suggest that this may not be the case for patients undergoing less extensive colorectal procedures, such as colorectal resections.

This study has some limitations. First, since we used secondary data from a cohort study, we were unable to formulate psychometric hypotheses prior to data collection, as recommended by COSMIN [97]. Additionally, our sample was limited to patients who met specific inclusion criteria and consented to participate in the primary study, which may restrict the generalizability of our findings. For instance, our cohort consisted of relatively healthy patients, with 64% having an ASA score of ≤ 2 . As a result, we cannot rule out the possibility that different results might be observed in more comorbid patients undergoing colorectal surgery. Finally, our study was conducted in institutions with established ERPs and a high rate of laparoscopic surgeries (78%), so our results may not be generalizable to other care contexts.

2.5 Conclusion

This study supports that PROMIS-29 has limited content validity, construct validity, and responsiveness for measuring recovery within 4 weeks post-colorectal surgery. The absence of robust PROMs for assessing recovery in this context remains an important research gap. There is a pressing need to develop psychometrically sound PROMs to measure postoperative recovery after colorectal surgery in accordance with existing recovery frameworks[46] and optimal methodological standards[34, 57, 97].

2.6 Disclosures

Francesca Fermi, Samin Shirzadi, Ghadeer Olleik, Makena Pook, Maxime Lapointe Gagner, Sarah Al Ben Ali, Philip Nguyen-Powanda, Elahe Khorasani, Tahereh Najafi Ghezeljeh, Naser Alali MD, Katy Dmowski, Pepa Kaneva, Nicolò Pecorelli, Liane S. Feldman, Marylise Boutros, Lawrence Lee, Julio F. Fiore Jr have no disclosures in relation to this study.

CHAPTER 3 - CONCLUSION AND FUTURE DIRECTIONS

The need for high-quality and psychometrically sound PROMs to measure recovery after surgery has been emphasized in recent literature[58]. The consensus statement by ASER/POQI identified PROMIS-29 as a potential tool to fill this gap [50]. As a generic multidimensional questionnaire, PROMIS-29 encompasses both physical and mental domains and was developed according to optimal scientific standards and modern psychometric methods (item response theory) [57]. Although this questionnaire showed promising results in measuring recovery after pancreatic surgery [91], evidence of its psychometric properties after colorectal surgery remains limited [58]. Findings from this thesis research supports that PROMIS-29 has limited content validity, construct validity, and responsiveness when measuring recovery after colorectal surgery. This suggests that, in this specific context, the PROMIS-29 does not meet the minimum quality standards required by ISOQOL [54]. Indeed, our findings support that PROMIS-29 only covers 36% of recovery-related themes important for the patients, it is unable to discriminate between patients with different recovery trajectories (e.g., with vs. without complications), and is only responsive to recovery trajectories up to two weeks post-surgery. A systematic review conducted by Fiore et al. in 2018 underlined the current absence of questionnaires supported by robust psychometric properties [58].

In line with the results, future studies should focus on developing a sound PROM to effectively evaluate postoperative recovery after colorectal and other abdominal surgeries. In alignment with FDA and ISPOR guidelines [57, 107, 108], an ongoing project by our group aims to bridge this research gap by developing a psychometrically robust PROM to measure recovery after abdominal surgery. The initial step, led by Alam et al. in 2019 [46], involved developing a conceptual framework of recovery post-abdominal surgery. Through interviews

with 30 patients across four countries, 39 essential recovery themes were identified. Based on this framework, an item bank with 59 questions was created. The subsequent phase, currently ongoing, entails the use of Rasch Measurement Theory to guide the selection of items and development of scoring rules in a cohort study including 100 patients undergoing abdominal procedures. Upon completion, the final steps will involve evaluating the measurement properties and developing electronic platforms to aid data collection and facilitate the use of computerized adaptive testing.

The development of a robust recovery-specific PROM will bring important benefits to research by providing a robust patient-centered outcome measure for comparative-effectiveness research of innovations in abdominal surgery. When used in clinical practice, this PROM will address the needs of patients, caregivers, and payers who seek information about the time required to recover “back to normal” for the safe resumption of leisure and work activities. Also, the collection of PRO data through mobile operating systems (e.g., mobile phones, tablets) and web portals will empower patients to track their own recovery trajectory in real-time and potentially identify complications early when they may be more easily treated.

In the future, we envision that recovery-specific PROM data will be seamlessly integrated into electronic health records, providing a unique opportunity for recovery auditing and database-driven research. This has the potential to streamline the identification of impactful interventions to improve recovery, enrich patient-centered surgical decision-making, and set recovery benchmarks [109]. By leveraging these advancements, we can transform surgical recovery into a more precise, patient-centered, and data-driven process, ultimately enhancing the overall quality of perioperative care.

REFERENCES

1. Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen Y-J, Ciombor KK, Cohen S, Cooper HS, Deming D, Engstrom PF, Garrido-Laguna I, Grem JL, Grothey A, Hochster HS, Hoffe S, Hunt S, Kamel A, Kirilcuk N, Krishnamurthi S, Messersmith WA, Meyerhardt J, Miller ED, Mulcahy MF, Murphy JD, Nurkin S, Saltz L, Sharma S, Shibata D, Skibber JM, Sofocleous CT, Stoffel EM, Stotsky-Himelfarb E, Willett CG, Wuthrick E, Gregory KM, Freedman-Cass DA (2018) NCCN Guidelines Insights: Colon Cancer, Version 2.2018. *Journal of the National Comprehensive Cancer Network* 16:359–369. <https://doi.org/10.6004/jnccn.2018.0021>
2. Benson AB, Venook AP, Al-Hawary MM, Arain MA, Chen Y-J, Ciombor KK, Cohen S, Cooper HS, Deming D, Garrido-Laguna I, Grem JL, Gunn A, Hoffe S, Hubbard J, Hunt S, Kirilcuk N, Krishnamurthi S, Messersmith WA, Meyerhardt J, Miller ED, Mulcahy MF, Nurkin S, Overman MJ, Parikh A, Patel H, Pedersen K, Saltz L, Schneider C, Shibata D, Skibber JM, Sofocleous CT, Stoffel EM, Stotsky-Himelfarb E, Willett CG, Johnson-Chilla A, Gurski LA (2020) NCCN Guidelines Insights: Rectal Cancer, Version 6.2020. *Journal of the National Comprehensive Cancer Network* 18:806–815. <https://doi.org/10.6004/jnccn.2020.0032>
3. Hall J, Hardiman K, Lee S, Lightner A, Stocchi L, Paquette IM, Steele SR, Feingold DL (2020) The American Society of Colon and Rectal Surgeons Clinical Practice Guidelines for the Treatment of Left-Sided Colonic Diverticulitis. *Dis Colon Rectum* 63:728–747. <https://doi.org/10.1097/DCR.0000000000001679>
4. You H, Sweeny A, Cooper ML, Von Papen M, Innes J (2019) The management of diverticulitis: a review of the guidelines. *Medical Journal of Australia* 211:421–427. <https://doi.org/10.5694/mja2.50276>
5. Spinelli A, Bonovas S, Burisch J, Kucharzik T, Adamina M, Annese V, Bachmann O, Bettenworth D, Chaparro M, Czuber-Dochan W, Eder P, Ellul P, Fidalgo C, Fiorino G, Gionchetti P, Gisbert JP, Gordon H, Hedin C, Holubar S, Iacucci M, Karmiris K, Katsanos K, Kopylov U, Lakatos PL, Lytras T,

- Lyutakov I, Noor N, Pellino G, Piovani D, Savarino E, Selvaggi F, Verstockt B, Doherty G, Raine T, Panis Y (2022) ECCO Guidelines on Therapeutics in Ulcerative Colitis: Surgical Treatment. *J Crohns Colitis* 16:179–189.
<https://doi.org/10.1093/ecco-jcc/jjab177>
6. Adamina M, Bonovas S, Raine T, Spinelli A, Warusavitarne J, Armuzzi A, Bachmann O, Bager P, Biancone L, Bokemeyer B, Bossuyt P, Burisch J, Collins P, Doherty G, El-Hussuna A, Ellul P, Fiorino G, Frei-Lanter C, Furfaro F, Gingert C, Gionchetti P, Gisbert JP, Gomollon F, González Lorenzo M, Gordon H, Hlavaty T, Juillerat P, Katsanos K, Kopylov U, Krustins E, Kucharzik T, Lytras T, Maaser C, Magro F, Marshall JK, Myrelid P, Pellino G, Rosa I, Sabino J, Savarino E, Stassen L, Torres J, Uzzan M, Vavricka S, Verstockt B, Zmora O (2020) ECCO Guidelines on Therapeutics in Crohn’s Disease: Surgical Treatment. *J Crohns Colitis* 14:155–168. <https://doi.org/10.1093/ecco-jcc/jjz187>
 7. Canadian Cancer Statistics 2023. Available at: <https://cancer.ca/en/cancer-information/resources/publications/canadian-cancer-statistics-2023>
 8. ACS releases colorectal cancer estimates for 2024. Available at: <https://colorectalcaner.org/article/acs-releases-colorectal-cancer-estimates-2024>
 9. Carli F (2015) Physiologic considerations of Enhanced Recovery After Surgery (ERAS) programs: implications of the stress response. *Canadian Journal of Anesthesia/Journal canadien d’anesthésie* 62:110–119.
<https://doi.org/10.1007/s12630-014-0264-0>
 10. Lee L, Tran T, Mayo NE, Carli F, Feldman LS (2014) What does it really mean to “recover” from an operation? *Surgery* 155:211–216.
<https://doi.org/10.1016/j.surg.2013.10.002>
 11. Feldman LS, Kaneva P, Demyttenaere S, Carli F, Fried GM, Mayo NE (2009) Validation of a physical activity questionnaire (CHAMPS) as an indicator of postoperative recovery after laparoscopic cholecystectomy. *Surgery* 146:31–39.
<https://doi.org/10.1016/j.surg.2009.02.019>
 12. Lawrence VA, Hazuda HP, Cornell JE, Pederson T, Bradshaw PT, Mulrow CD, Page CP (2004) Functional independence after major abdominal surgery in the elderly. *J Am Coll Surg* 199:762–772.
<https://doi.org/10.1016/j.jamcollsurg.2004.05.280>

13. Collins TC, Daley J, Henderson WH, Khuri SF (1999) Risk Factors for Prolonged Length of Stay After Major Elective Surgery. *Ann Surg* 230:251. <https://doi.org/10.1097/00000658-199908000-00016>
14. Scarborough JE, Schumacher J, Kent KC, Heise CP, Greenberg CC (2017) Associations of Specific Postoperative Complications With Outcomes After Elective Colon Resection. *JAMA Surg* 152:e164681. <https://doi.org/10.1001/jamasurg.2016.4681>
15. Chapuis PH, Bokey L, Keshava A, Rickard MJFX, Stewart P, Young CJ, Dent OF (2013) Risk Factors for Prolonged Ileus After Resection of Colorectal Cancer. *Ann Surg* 257:909–915. <https://doi.org/10.1097/SLA.0b013e318268a693>
16. Trencheva K, Morrissey KP, Wells M, Mancuso CA, Lee SW, Sonoda T, Michelassi F, Charlson ME, Milsom JW (2013) Identifying Important Predictors for Anastomotic Leak After Colon and Rectal Resection. *Ann Surg* 257:108–113. <https://doi.org/10.1097/SLA.0b013e318262a6cd>
17. Wood T, Aarts M-A, Okrainec A, Pearsall E, Victor JC, McKenzie M, Rotstein O, McLeod RS, Forbes S, Tuszynska H, Onorato L, Macdonald J, Broomfield D, Caycedo A, Hick T, Clement G, Blakely S, MacDonald H, Van Vlymen J, McPherson N, Scott E, Schlachta C, Harle C, Stillwell K, Chambers A, Govindarajan A, Siddiqui N, Sarvi E, Memon A, Stotland P, Irwin D, Wong M, Vanta L, Fenech D, Morningstar B, Avila E, Speke A, Rotstein O, Hamdy A, Park J, Reddigan J, Cadeddu M, Horner C, Dunn M, Adeyemo T, Lindsay D, Savoie M, Tomat L, Auer R, Szeto M, McGrath M, Sadiq M, Sinclair J, Mapeso G, Bonneville S, Olson D, Woodbeck E, Aarts M-A, Saby C, Sukovic T, O'Brien C, McCluskey S, Zamora T, Okrainec A, Cheung F, Lopez C, Mirkolaei AM (2018) Emergency Room Visits and Readmissions Following Implementation of an Enhanced Recovery After Surgery (iERAS) Program. *Journal of Gastrointestinal Surgery* 22:259–266. <https://doi.org/10.1007/s11605-017-3555-2>
18. Huisman DE, Ingwersen EW, Luttikhoud J, Slooter GD, Kazemier G, Daams F (2024) Prediction of Length of Stay After Colorectal Surgery Using

- Intraoperative Risk Factors. *Annals of Surgery Open* 5:e478.
<https://doi.org/10.1097/AS9.0000000000000478>
19. Buitrago-Ruiz M, Martinez-Nicolas I, Soria-Aledo V (2023) Validation of prolonged length of stay as a reliable measure of failure to rescue in colorectal surgery. *Asian J Surg* 46:126–131. <https://doi.org/10.1016/j.asjsur.2022.02.003>
 20. Vonlanthen R, Slankamenac K, Breitenstein S, Puhon MA, Muller MK, Hahnloser D, Hauri D, Graf R, Clavien P-A (2011) The Impact of Complications on Costs of Major Surgical Procedures. *Ann Surg* 254:907–913. <https://doi.org/10.1097/SLA.0b013e31821d4a43>
 21. Downey CL, Bainbridge J, Jayne DG, Meads DM (2023) Impact of in-hospital postoperative complications on quality of life up to 12 months after major abdominal surgery. *British Journal of Surgery* 110:1206–1212. <https://doi.org/10.1093/bjs/znad167>
 22. Bowyer A, Royse C (2016) The importance of postoperative quality of recovery: influences, assessment, and clinical and prognostic implications. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie* 63:176–183. <https://doi.org/10.1007/s12630-015-0508-7>
 23. Lee L, Feldman LS (2017) Improving Surgical Value and Culture Through Enhanced Recovery Programs. *JAMA Surg* 152:299. <https://doi.org/10.1001/jamasurg.2016.5056>
 24. Kehlet H (1997) Multimodal approach to control postoperative pathophysiology and rehabilitation. *Br J Anaesth* 78:606–617. <https://doi.org/10.1093/bja/78.5.606>
 25. Gustafsson UO, Scott MJ, Hubner M, Nygren J, Demartines N, Francis N, Rockall TA, Young-Fadok TM, Hill AG, Soop M, de Boer HD, Urman RD, Chang GJ, Fichera A, Kessler H, Grass F, Whang EE, Fawcett WJ, Carli F, Lobo DN, Rollins KE, Balfour A, Baldini G, Riedel B, Ljungqvist O (2019) Guidelines for Perioperative Care in Elective Colorectal Surgery: Enhanced Recovery After Surgery (ERAS[®]) Society Recommendations: 2018. *World J Surg* 43:659–695. <https://doi.org/10.1007/s00268-018-4844-y>

26. Irani JL, Hedrick TL, Miller TE, Lee L, Steinhagen E, Shogan BD, Goldberg JE, Feingold DL, Lightner AL, Paquette IM (2023) Clinical Practice Guidelines for Enhanced Recovery After Colon and Rectal Surgery From the American Society of Colon and Rectal Surgeons and the Society of American Gastrointestinal and Endoscopic Surgeons. *Dis Colon Rectum* 66:15–40.
<https://doi.org/10.1097/DCR.0000000000002650>
27. Schwenk W, Haase O, Neudecker JJ, Müller JM (2005) Short term benefits for laparoscopic colorectal resection. *Cochrane Database of Systematic Reviews* 2008:. <https://doi.org/10.1002/14651858.CD003145.pub2>
28. Zhuang C-L, Ye X-Z, Zhang X-D, Chen B-C, Yu Z (2013) Enhanced Recovery After Surgery Programs Versus Traditional Care for Colorectal Surgery. *Dis Colon Rectum* 56:667–678. <https://doi.org/10.1097/DCR.0b013e3182812842>
29. Kehlet H, Wilmore DW (2008) Evidence-Based Surgical Care and the Evolution of Fast-Track Surgery. *Ann Surg* 248:189–198.
<https://doi.org/10.1097/SLA.0b013e31817f2c1a>
30. Nicholson A, Lowe MC, Parker J, Lewis SR, Alderson P, Smith AF (2014) Systematic review and meta-analysis of enhanced recovery programmes in surgical patients. *British Journal of Surgery* 101:172–188.
<https://doi.org/10.1002/bjs.9394>
31. Zhuang C-L, Huang D-D, Chen F-F, Zhou C-J, Zheng B-S, Chen B-C, Shen X, Yu Z (2015) Laparoscopic versus open colorectal surgery within enhanced recovery after surgery programs: a systematic review and meta-analysis of randomized controlled trials. *Surg Endosc* 29:2091–2100.
<https://doi.org/10.1007/s00464-014-3922-y>
32. Neville A, Lee L, Antonescu I, Mayo NE, Vassiliou MC, Fried GM, Feldman LS (2014) Systematic review of outcomes used to evaluate enhanced recovery after surgery. *British Journal of Surgery* 101:159–170.
<https://doi.org/10.1002/bjs.9324>
33. Antonescu I, Mueller CL, Fried GM, Vassiliou MC, Mayo NE, Feldman LS (2014) Outcomes reported in high-impact surgical journals. *British Journal of Surgery* 101:582–589. <https://doi.org/10.1002/bjs.9450>

34. Walton MK, Powers JH, Hobart J, Patrick D, Marquis P, Vamvakas S, Isaac M, Molsen E, Cano S, Burke LB (2015) Clinical Outcome Assessments: Conceptual Foundation—Report of the ISPOR Clinical Outcomes Assessment – Emerging Good Practices for Outcomes Research Task Force. *Value in Health* 18:741–752. <https://doi.org/10.1016/j.jval.2015.08.006>
35. Almoudaris AM, Clark S, Vincent C, Faiz O (2011) Establishing quality in colorectal surgery. *Colorectal Disease* 13:961–973. <https://doi.org/10.1111/j.1463-1318.2010.02355.x>
36. Lee L, Mata J, Ghitulescu GA, Boutros M, Charlebois P, Stein B, Liberman AS, Fried GM, Morin N, Carli F, Latimer E, Feldman LS (2015) Cost-effectiveness of Enhanced Recovery Versus Conventional Perioperative Management for Colorectal Surgery. *Ann Surg* 262:1026–1033. <https://doi.org/10.1097/SLA.0000000000001019>
37. Benson RT, Drew JC, Galland RB (2006) A waiting list to go home: an analysis of delayed discharges from surgical beds. *Ann R Coll Surg Engl* 88:650–2. <https://doi.org/10.1308/003588406X149246>
38. Borza T, Oerline MK, Skolarus TA, Norton EC, Ryan AM, Ellimoottil C, Dimick JB, Shahinian VB, Hollenbeck BK (2018) Association of the Hospital Readmissions Reduction Program With Surgical Readmissions. *JAMA Surg* 153:243. <https://doi.org/10.1001/jamasurg.2017.4585>
39. Maessen J, Dejong CHC, Hausel J, Nygren J, Lassen K, Andersen J, Kessels AGH, Revhaug A, Kehlet H, Ljungqvist O, Fearon KCH, von Meyenfeldt MF (2007) A protocol is not enough to implement an enhanced recovery programme for colorectal resection. *British Journal of Surgery* 94:224–231. <https://doi.org/10.1002/bjs.5468>
40. Balvardi S, Pecorelli N, Castelino T, Niculiseanu P, Liberman AS, Charlebois P, Stein B, Carli F, Mayo NE, Feldman LS, Fiore JF (2018) Measuring In-Hospital Recovery After Colorectal Surgery Within a Well-Established Enhanced Recovery Pathway: A Comparison Between Hospital Length of Stay and Time to Readiness for Discharge. *Dis Colon Rectum* 61:854–860. <https://doi.org/10.1097/DCR.0000000000001061>

41. Martin RCG, Brennan MF, Jaques DP (2002) Quality of Complication Reporting in the Surgical Literature. *Ann Surg* 235:803–813.
<https://doi.org/10.1097/00000658-200206000-00007>
42. Slankamenac K, Graf R, Barkun J, Puhan MA, Clavien P-A (2013) The Comprehensive Complication Index. *Ann Surg* 258:1–7.
<https://doi.org/10.1097/SLA.0b013e318296c732>
43. Dindo D, Demartines N, Clavien P-A (2004) Classification of Surgical Complications. *Ann Surg* 240:205–213.
<https://doi.org/10.1097/01.sla.0000133083.54934.ae>
44. Allvin R, Berg K, Idvall E, Nilsson U (2007) Postoperative recovery: a concept analysis. *J Adv Nurs* 57:552–558. <https://doi.org/10.1111/j.1365-2648.2006.04156.x>
45. Rajabiyazdi F, Alam R, Pal A, Montanez J, Law S, Pecorelli N, Watanabe Y, Chiavegato LD, Falconi M, Hirano S, Mayo NE, Lee L, Feldman LS, Fiore JF (2021) Understanding the Meaning of Recovery to Patients Undergoing Abdominal Surgery. *JAMA Surg* 156:758.
<https://doi.org/10.1001/jamasurg.2021.1557>
46. Alam R, Montanez J, Law S, Lee L, Pecorelli N, Watanabe Y, Chiavegato LD, Falconi M, Hirano S, Mayo NE, Feldman LS, Fiore JF (2020) Development of a conceptual framework of recovery after abdominal surgery. *Surg Endosc* 34:2665–2674. <https://doi.org/10.1007/s00464-019-07044-x>
47. Meadows KA (2011) Patient-reported outcome measures: an overview. *Br J Community Nurs* 16:146–151. <https://doi.org/10.12968/bjcn.2011.16.3.146>
48. Black N (2013) Patient reported outcome measures could help transform healthcare. *BMJ* 346:f167–f167. <https://doi.org/10.1136/bmj.f167>
49. Weldring T, Smith SMS (2013) Article Commentary: Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Serv Insights* 6:HSI.S11093. <https://doi.org/10.4137/HSI.S11093>
50. Abola RE, Bennett-Guerrero E, Kent ML, Feldman LS, Fiore JF, Shaw AD, Thacker JKM, Gan TJ, Miller TE, Hedrick TL, McEvoy MD, Mythen MG, Bergamaschi R, Gupta R, Holubar SD, Senagore AJ, Wischmeyer PE, Carli F,

- Evans DC, Guilbert S, Kozar R, Pryor A, Thiele RH, Everett S, Grocott M (2018) American Society for Enhanced Recovery and Perioperative Quality Initiative Joint Consensus Statement on Patient-Reported Outcomes in an Enhanced Recovery Pathway. *Anesth Analg* 126:1874–1882
51. Bartels S AL (2010) Quality of life after laparoscopic and open colorectal surgery: A systematic review. *World J Gastroenterol* 16:5035. <https://doi.org/10.3748/wjg.v16.i40.5035>
 52. Li D, Jensen C (2019) Patient Satisfaction and Quality of Life with Enhanced Recovery Protocols. *Clin Colon Rectal Surg* 32:138–144. <https://doi.org/10.1055/s-0038-1676480>
 53. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research* 19:539–549. <https://doi.org/10.1007/s11136-010-9606-8>
 54. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, Schwartz C, Revicki DA, Moinpour CM, McLeod LD, Lyons JC, Lenderking WR, Hinds PS, Hays RD, Greenhalgh J, Gershon R, Feeny D, Fayers PM, Cella D, Brundage M, Ahmed S, Aaronson NK, Butt Z (2013) ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research* 22:1889–1905. <https://doi.org/10.1007/s11136-012-0344-y>
 55. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW (2010) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 63:737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
 56. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Mokkink LB (2018) COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research* 27:1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>

57. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. [http://www.fda.gov/downloads/Drugs/ Guidances/UCM193282.pdf](http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf). Published 2009. Accessed 23 Sept
58. Fiore JF, Figueiredo S, Balvardi S, Lee L, Nauche B, Landry T, Mayo NE, Feldman LS (2018) How Do We Value Postoperative Recovery? *Ann Surg* 267:656–669. <https://doi.org/10.1097/SLA.0000000000002415>
59. Intro to PROMIS. Available at:[https://www.healthmeasures.net/explore-measurement-systems/ promis/intro-to- promis](https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis). Accessed 27 May 2024.
60. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M (2007) The Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care* 45:S3–S11. <https://doi.org/10.1097/01.mlr.0000258615.42478.55>
61. Differences Between PROMIS Measures. Available at: <https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/differences-between-promis-measures>. Accessed 27 May 2024
62. Slides Introducing HealthMeasures - PROMIS. Available at: <https://www.healthmeasures.net/resource-center/measurement-science/slides-introducing-healthmeasures>. Accessed 27 may 2024.
63. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, DeVellis R, DeWalt D, Fries JF, Gershon R, Hahn EA, Lai J-S, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R (2010) The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 63:1179–1194. <https://doi.org/10.1016/j.jclinepi.2010.04.011>
64. Amtmann D, Cook KF, Jensen MP, Chen W-H, Choi S, Revicki D, Cella D, Rothrock N, Keefe F, Callahan L, Lai J-S (2010) Development of a PROMIS item bank to measure pain interference. *Pain* 150:173–182. <https://doi.org/10.1016/j.pain.2010.04.025>

65. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Lai J, Cella D (2007) Psychometric Evaluation and Calibration of Health-Related Quality of Life Item Banks. *Med Care* 45:S22–S31.
<https://doi.org/10.1097/01.mlr.0000250483.85507.04>
66. Item Response Theory (IRT).
https://www.healthmeasures.net/index.php?option=com_content&view=category&layout=blog&id=134&Itemid=938. Accessed 27 May 2024.
67. Computer Adaptive Tests (CATs). Available at:
https://www.healthmeasures.net/index.php?option=com_content&view=category&layout=blog&id=164&Itemid=1133. Accessed 27 May 2024.
68. Liu H, Cella D, Gershon R, Shen J, Morales LS, Riley W, Hays RD (2010) Representativeness of the Patient-Reported Outcomes Measurement Information System Internet panel. *J Clin Epidemiol* 63:1169–1178.
<https://doi.org/10.1016/j.jclinepi.2009.11.021>
69. Interpret Scores. Available at: <https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis>. Accessed 27 May 2024.
70. PROMIS-29 v2.1. Available at:
https://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manuals/_PROMIS_Adult_Profile_Scoring_Manual.pdf. Accessed 27 May 2024.
71. Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, Johnston KL, Shablesky-Cade MA, Pilkonis PA (2010) Development and Validation of Patient-Reported Outcome Measures for Sleep Disturbance and Sleep-Related Impairments. *Sleep* 33:781–792. <https://doi.org/10.1093/sleep/33.6.781>
72. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D (2011) Item Banks for Measuring Emotional Distress From the Patient-Reported Outcomes Measurement Information System (PROMIS®): Depression, Anxiety, and Anger. *Assessment* 18:263–283. <https://doi.org/10.1177/1073191111411667>
73. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE (2014) The PROMIS Physical Function item bank was calibrated to a standardized metric

- and shown to improve measurement efficiency. *J Clin Epidemiol* 67:516–526.
<https://doi.org/10.1016/j.jclinepi.2013.10.024>
74. Craig BM, Reeve BB, Brown PM, Cella D, Hays RD, Lipscomb J, Simon Pickard A, Revicki DA (2014) US Valuation of Health Outcomes Measured Using the PROMIS-29. *Value in Health* 17:846–853.
<https://doi.org/10.1016/j.jval.2014.09.005>
 75. Hays RD, Spritzer KL, Schalet BD, Cella D (2018) PROMIS®-29 v2.0 profile physical and mental health summary scores. *Quality of Life Research* 27:1885–1891. <https://doi.org/10.1007/s11136-018-1842-3>
 76. Kang CY, Halabi WJ, Luo R, Pigazzi A, Nguyen NT, Stamos MJ (2012) Laparoscopic Colorectal Surgery. *Archives of Surgery* 147:..
<https://doi.org/10.1001/archsurg.2012.358>
 77. Schilling PL, Dimick JB, Birkmeyer JD (2008) Prioritizing Quality Improvement in General Surgery. *J Am Coll Surg* 207:698–704.
<https://doi.org/10.1016/j.jamcollsurg.2008.06.138>
 78. Arriaga AF, Lancaster RT, Berry WR, Regenbogen SE, Lipsitz SR, Kaafarani HMA, Elbardissi AW, Desai P, Ferzoco SJ, Bleday R, Breen E, Kastrinakis W V., Rubin MS, Gawande AA (2009) The Better Colectomy Project. *Ann Surg* 250:507–513. <https://doi.org/10.1097/SLA.0b013e3181b672bc>
 79. Ghignone F, Hernandez P, Mahmoud NN, Ugolini G (2020) Functional recovery in senior adults undergoing surgery for colorectal cancer: Assessment tools and strategies to preserve functional status. *European Journal of Surgical Oncology* 46:387–393. <https://doi.org/10.1016/j.ejso.2020.01.003>
 80. Lee L, Dumitra T, Fiore JF, Mayo NE, Feldman LS (2015) How well are we measuring postoperative “recovery” after abdominal surgery? *Quality of Life Research* 24:2583–2590. <https://doi.org/10.1007/s11136-015-1008-5>
 81. Feldman LS, Lee L, Fiore J (2015) What outcomes are important in the assessment of Enhanced Recovery After Surgery (ERAS) pathways? *Canadian Journal of Anesthesia/Journal canadien d’anesthésie* 62:120–130.
<https://doi.org/10.1007/s12630-014-0263-1>

82. Alonso J, Bartlett SJ, Rose M, Aaronson NK, Chaplin JE, Efficace F, Leplège A, LU A, Tulskey DS, Raat H, Ravens-Sieberer U, Revicki D, Terwee CB, Valderas JM, Cella D, Forrest CB (2013) The case for an international patient-reported outcomes measurement information system (PROMIS®) initiative. *Health Qual Life Outcomes* 11:210. <https://doi.org/10.1186/1477-7525-11-210>
83. Fries JF, Bruce B, Cella D (2005) The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol* 23:S53-7
84. COSMIN. Available at: <https://www.cosmin.nl/>. Accessed 27 May 2024.
85. Olleik G, Lapointe-Gagner M, Jain S, Shirzadi S, Nguyen-Powanda P, Al Ben Ali S, Ghezeljeh TN, Elhaj H, Alali N, Fermi F, Pook M, Mousoulis C, Almusailleem A, Farag N, Dmowski K, Cutler D, Kaneva P, Agnihotram R V., Feldman LS, Boutros M, Lee L, Fiore JF (2024) Opioid use patterns following discharge from elective colorectal surgery: a prospective cohort study. *Surg Endosc*. <https://doi.org/10.1007/s00464-024-11322-8>
86. Lee L, Eustache J, Baldini G, Liberman AS, Charlebois P, Stein B, Fiore JF, Feldman LS (2022) Enhanced Recovery 2.0 – Same Day Discharge With Mobile App Follow-up After Minimally Invasive Colorectal Surgery. *Ann Surg* 276:e812–e818. <https://doi.org/10.1097/SLA.0000000000004962>
87. Scoring Instruction. Available at: <https://www.healthmeasures.net/score-and-interpret/calculate-scores/scoring-instructions>. Accessed 27 May 2024.
88. World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. Geneva: WHO.
89. Tucker CA, Escorpizo R, Cieza A, Lai JS, Stucki G, Ustun TB, Kostanjsek N, Cella D, Forrest CB (2014) Mapping the content of the Patient-Reported Outcomes Measurement Information System (PROMIS®) using the International Classification of Functioning, Health and Disability. *Quality of Life Research* 23:2431–2438. <https://doi.org/10.1007/s11136-014-0691-y>
90. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW (2007) Quality criteria were proposed for

- measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
91. Pecorelli N, Guarneri G, Vallorani A, Limongi C, Licinio AW, Di Salvo F, Tamburrino D, Partelli S, Crippa S, Falconi M (2023) Validation of the PROMIS-29 Questionnaire as a Measure of Recovery After Pancreatic Surgery. *Ann Surg* 278:732–739. <https://doi.org/10.1097/SLA.0000000000006020>
 92. Balvardi S, Pecorelli N, Castelino T, Niculiseanu P, Liberman AS, Charlebois P, Stein B, Carli F, Mayo NE, Feldman LS, Fiore JF (2019) Construct Validity and Responsiveness of the Abdominal Surgery Impact Scale in the Context of Recovery After Colorectal Surgery. *Dis Colon Rectum* 62:309–317. <https://doi.org/10.1097/DCR.0000000000001288>
 93. Lee L, Mata J, Augustin BR, Carli F, Morin N, Latimer E, Feldman LS (2014) A comparison of the validity of two indirect utility instruments as measures of postoperative recovery. *Journal of Surgical Research* 190:79–86. <https://doi.org/10.1016/j.jss.2014.02.016>
 94. Dowson HM, Ballard K, Gage H, Jackson D, Williams P, Rockall TA (2013) Quality of Life in the First 6 Weeks Following Laparoscopic and Open Colorectal Surgery. *Value in Health* 16:367–372. <https://doi.org/10.1016/j.jval.2012.11.005>
 95. Fucini C, Gattai R, Urena C, Bandettini L, Elbetti C (2008) Quality of Life Among Five-Year Survivors After Treatment for Very Low Rectal Cancer With or Without a Permanent Abdominal Stoma. *Ann Surg Oncol* 15:1099–1106. <https://doi.org/10.1245/s10434-007-9748-2>
 96. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, De Vet HCW (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research* 19:539–549. <https://doi.org/10.1007/s11136-010-9606-8>
 97. Mokkink Cecilia AC Prinsen Donald L Patrick Jordi Alonso Lex M Bouter, L. B. & Mokkink, C. L. COSMIN Study Design Checklist for Patient-Reported Outcome Measurement Instruments. www.cosmin.nl.

98. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
99. Tavakol M, Dennick R (2011) Making sense of Cronbach's alpha. *Int J Med Educ* 2:53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
100. Kleinke, K. (2017). Multiple imputation under violated distributional assumptions: A systematic evaluation of the assumed robustness of predictive mean matching. *Journal of Educational and Behavioral Statistics*, 42(4), 371–404
101. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 338:b2393–b2393. <https://doi.org/10.1136/bmj.b2393>
102. Thiese MS, Ronna B, Ott U (2016) P value interpretations and considerations. *J Thorac Dis* 8:E928–E931. <https://doi.org/10.21037/jtd.2016.08.16>
103. Terwee CB, Peipert JD, Chapman R, Lai J-S, Terluin B, Cella D, Griffiths P, Mokkink LB (2021) Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Quality of Life Research* 30:2729–2754. <https://doi.org/10.1007/s11136-021-02925-y>
104. Meaningful Change for PROMIS. Available at: <https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/meaningful-change>. Accessed 27 May 2024.
105. Mouelhi Y, Jouve E, Castelli C, Gentile S (2020) How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health Qual Life Outcomes* 18:136. <https://doi.org/10.1186/s12955-020-01344-w>
106. Rosengaard LO, Andersen MZ, Rosenberg J, Fonnes S (2024) Five aspects of research waste in biomedicine: A scoping review. *J Evid Based Med* 17:351–359. <https://doi.org/10.1111/jebm.12616>

107. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, Ring L (2011) Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting Concepts for a New PRO Instrument. *Value in Health* 14:967–977. <https://doi.org/10.1016/j.jval.2011.06.014>
108. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, Ring L (2011) Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2—Assessing Respondent Understanding. *Value in Health* 14:978–988. <https://doi.org/10.1016/j.jval.2011.06.013>
109. Pezold ML, Pusic AL, Cohen WA, Hollenberg JP, Butt Z, Flum DR, Temple LK (2016) Defining a Research Agenda for Patient-Reported Outcomes in Surgery. *JAMA Surg* 151:930. <https://doi.org/10.1001/jamasurg.2016.1640>
110. Guillou PJ, Quirke P, Thorpe H, Walker J, Jayne DG, Smith AM, Heath RM, Brown JM (2005) Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *The Lancet* 365:1718–1726. [https://doi.org/10.1016/S0140-6736\(05\)66545-2](https://doi.org/10.1016/S0140-6736(05)66545-2)
111. Buzby G, Knox L, Crosby L, Eisenberg J, Haakenson C, McNeal G, Page C, Reinhardt G, Williford W (1988) Study protocol: a randomized clinical trial of total parenteral nutrition in malnourished surgical patients. *Am J Clin Nutr* 47:366–381. <https://doi.org/10.1093/ajcn/47.2.366>
112. (2005) Guidelines for the Management of Adults with Hospital-acquired, Ventilator-associated, and Healthcare-associated Pneumonia. *Am J Respir Crit Care Med* 171:388–416. <https://doi.org/10.1164/rccm.200405-644ST>
113. Bellomo R (2005) Defining, Quantifying, and Classifying Acute Renal Failure. *Crit Care Clin* 21:223–237. <https://doi.org/10.1016/j.ccc.2004.12.001>
114. Dudeck MA, Horan TC, Peterson KD, Allen-Bridson K, Morrell GC, Pollock DA, Edwards JR (2011) National Healthcare Safety Network (NHSN) report,

- data summary for 2009, device-associated module. *Am J Infect Control* 39:349–367. <https://doi.org/10.1016/j.ajic.2011.04.011>
115. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG (1992) CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol* 13:606–8
 116. Rahbari NN, Weitz J, Hohenberger W, Heald RJ, Moran B, Ulrich A, Holm T, Wong WD, Tiet E, Moriya Y, Laurberg S, den Dulk M, van de Velde C, Büchler MW (2010) Definition and grading of anastomotic leakage following anterior resection of the rectum: A proposal by the International Study Group of Rectal Cancer. *Surgery* 147:339–351. <https://doi.org/10.1016/j.surg.2009.10.012>
 117. Lång M, Niskanen M, Miettinen P, Alhava E, Takala J (2002) Outcome and resource utilization in gastroenterological surgery. *British Journal of Surgery* 88:1006–1014. <https://doi.org/10.1046/j.0007-1323.2001.01812.x>

APPENDIX

Supplemental Digital Content 1. Missing PROMIS-29 data at different time points (imputed)

	Baseline	POW1	POW2	POW3	POW4
<i>PROMIS-29 domains</i>					
- Physical Function	6 (2)	30 (11)	32 (11)	38 (13)	25 (9)
- Anxiety	6 (2)	30 (11)	32 (11)	38 (13)	25 (9)
- Depression	6 (2)	30 (11)	32 (11)	38 (13)	25 (9)
- Fatigue	6 (2)	30 (11)	32 (11)	38 (13)	25 (9)
- Sleep disorders	7 (2)	30 (11)	32 (11)	39 (14)	26 (9)
- Ability to participate in social roles and activities	7 (2)	30 (11)	32 (11)	39 (14)	25 (9)
- Pain interference	6 (2)	30 (11)	32 (11)	39 (14)	25 (9)
- Pain intensity	6 (2)	30 (11)	32 (11)	38 (13)	25 (9)
<i>PROMIS-29 summary scores</i>					
- PHS score	8 (3)	30 (11)	32 (11)	39 (14)	26 (9)
- MHS score	8 (3)	30 (11)	32 (11)	39 (14)	26 (9)

Data are expressed as n (%).

PHS, physical health summary score; *MHS*, mental health summary score; *POW*, postoperative week.

Supplemental Digital Content 2. 30-day postoperative complications (total n=97/282)

Type of complication	Definition	Frequency n (%)
INTRAOPERATIVE [110]		
<i>Clinically significant hemorrhage</i>	Intraoperative bleeding requiring transfusion of packed red blood cells (PRBC) during surgery or within 24 hours after surgery.	2 (1)
<i>Urinary tract injury</i>	Injury of the ureter or bladder requiring intraoperative repair.	1 (0)
<i>Cardiac or respiratory complications</i>	Any cardiovascular (e.g. cardiac arrhythmia, myocardial infarction) or respiratory (e.g. pneumothorax) complication occurring during surgery.	1 (0)
POSTOPERATIVE		
MEDICAL		
Cardiovascular		
<i>Cardiac arrhythmia</i>	ECG diagnosis of new arrhythmia requiring at least a pharmacologic intervention.[111]	2 (1)
<i>Deep vein thrombosis</i>	Radiological confirmation of deep vein thrombosis or anticoagulation started due to clinical findings.	1 (0)
<i>Pulmonary embolism</i>	Radiological evidence of pulmonary embolism.	2 (1)
Respiratory		
<i>Pneumoniae</i>	Hospital acquired pneumonia, defined as presence of lung infiltrate at chest x-ray accompanied with signs of infection and initiation of antibiotic treatment.[112]	2 (1)
<i>Lobar atelectasis</i>	Radiological finding of at least one lobar collapse.[111]	2 (1)
Other medical		
<i>Acute Kidney Injury</i>	Increase in serum creatinine $\times 2$ from baseline or reduction of glomerular filtration rate greater than 50%.[113]	5 (2)
<i>Urinary retention</i>	Reinsertion of indwelling urinary catheter after removal attempt or patient discharged with urinary drainage (excluding patients with permanent indwelling urinary catheter).	11 (4)
<i>Anemia</i>	Low serum hemoglobin requiring transfusion of PRBC, unrelated to any identified source of bleeding.	4 (1)
<i>Other gastrointestinal complications</i>	Any other complication of the gastrointestinal tract requiring treatment (e.g. blood per rectum, diarrhea, high stoma output).	15 (5)
<i>Psychiatric complications</i>	New psychiatric symptoms including delirium and depression, requiring pharmacological treatment.	2 (1)
Infectious		

<i>Urinary tract infection</i>	Upper or lower urinary symptoms and urine culture with no more than two species of organisms, at least one of which is a bacteria of $\geq 10^5$ CFU/ml.[114]	8 (3)
<i>Wound infection</i>	Purulent drainage, with or without positive culture, from the superficial incision or any sign or symptom of infection (e.g. pain or tenderness, localized swelling, redness) and superficial incision is deliberately opened by the surgeon or attending physician. Not included if part of intra-peritoneal abscess.[115]	9 (3)
<i>Intra- or retroperitoneal abscess</i>	Radiologic finding of deep collection of pus associated with systemic signs of infection or finding during reoperation.	10 (4)
<i>Other infectious complications</i>	Any other documented infectious complication (e.g. Clostridium difficile colitis).	5 (2)
<i>Surgical</i>		
<i>Anastomotic leak</i>	Documentation at reoperation OR documentation by imaging technique (e.g. radiologically, endoscopically) of leakage from the surgical connection between the two bowel ends into the abdomen or pelvis with either spillage and/or fluid collection around the anastomotic site or extravasation through a wound, drain site, or anus.[16] In the case of rectal surgery, a pelvic abscess close to the anastomosis is also considered as anastomotic leakage.[116]	4 (1)
<i>Mechanical bowel obstruction</i>	Documentation at reoperation OR radiologically of mechanical small or large bowel obstruction.	1 (0)
<i>Wound dehiscence</i>	Separation of the abdominal wall muscle fascia large enough to necessitate operative closure of the wound OR incisional hernia diagnosed after primary discharge. [117]	7 (2)
<i>Bleeding</i>	Any postoperative bleeding (e.g. intra-abdominal, gastrointestinal) requiring transfusion of at least 2 PRBC after surgery.	6 (2)
<i>Ileus (primary postoperative ileus)</i>	Abdominal distention OR vomiting associated with intolerance of solid food intake or inability to pass gas or stool beyond POD3 (target day for discharge), unrelated to any other ongoing complication.	26 (9)
<i>Other surgical complications</i>	Any other surgical complication necessitating treatment or delaying discharge (e.g. abdominal wall hematoma).	3 (1)
<i>Anesthesia-related</i>		
<i>Anesthesia-related complications</i>	Any anesthesia-related complication occurring after surgery (e.g. peripheral nerve injuries).	2 (1)
<i>Symptoms delaying discharge</i>		
<i>Pain</i>	Uncontrolled pain requiring prolonged treatment delaying discharge, unrelated to any other complication.	2 (1)

The data are reported as a frequency of events among the entire population. Overall, 9% (n=24/282) of patients experienced more than one complication.

-
- ¹ Guillo PJ, Quirke P, Thorpe H, et al. Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *The Lancet*. 2005;365:1718–1726.
- ² Buzby G, Knox L, Crosby L, et al. Study protocol: a randomized clinical trial of total parenteral nutrition in malnourished surgical patients. *Am J Clin Nutr*. 1988;47:366–381.
- ³ Guidelines for the Management of Adults with Hospital-acquired, Ventilator-associated, and Healthcare-associated Pneumonia. *Am J Respir Crit Care Med*. 2005;171:388–416.
- ⁴ Bellomo R. Defining, Quantifying, and Classifying Acute Renal Failure. *Crit Care Clin*. 2005;21:223–237.
- ⁵ Dudeck MA, Horan TC, Peterson KD, et al. National Healthcare Safety Network (NHSN) report, data summary for 2009, device-associated module. *Am J Infect Control*. 2011;39:349–367.
- ⁶ Horan TC, Gaynes RP, Martone WJ, et al. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol*. 1992;13:606–8.
- ⁷ Trencheva K, Morrissey KP, Wells M, et al. Identifying Important Predictors for Anastomotic Leak After Colon and Rectal Resection. *Ann Surg*. 2013;257:108–113.
- ⁸ Rahbari NN, Weitz J, Hohenberger W, et al. Definition and grading of anastomotic leakage following anterior resection of the rectum: A proposal by the International Study Group of Rectal Cancer. *Surgery*. 2010;147:339–351.
- ⁹ Lång M, Niskanen M, Miettinen P, et al. Outcome and resource utilization in gastroenterological surgery. *British Journal of Surgery*. 2002;88:1006–1014.

Supplemental Digital Content 3. Domain-specific PROMIS-29 scores, internal consistency, and floor and ceiling effects.

PROMIS-29 domains	Sample (n)	Mean $\pm$ SD	Median [IQR]	Cronbach Alpha	Floor Effect (%)	Ceiling Effect (%)
<i>Physical function (range 22.5-57)</i>						
Preoperatively	276	49 $\pm$ 9	57 [43-57]	0.89	1	53
POW1	252	40 $\pm$ 8	39 [34-43]	0.81	2	8
POW2	250	43 $\pm$ 8	42 [38-48]	0.83	1	16
POW3	244	46 $\pm$ 8	45 [40-57]	0.86	1	26
POW4	257	48 $\pm$ 9	48 [42-57]	0.87	1	40
<i>Anxiety (range 40.3 – 81.6)</i>						
Preoperatively	276	56 $\pm$ 9	56 [51-61]	0.90	16	1
POW1	252	52 $\pm$ 9	54 [40-58]	0.88	26	1
POW2	250	50 $\pm$ 9	51 [40-56]	0.90	38	0
POW3	244	49 $\pm$ 9	48 [40-56]	0.92	44	0
POW4	257	48 $\pm$ 9	48 [40-56]	0.91	49	0
<i>Depression (range 41 – 79.4)</i>						
Preoperatively	276	49 $\pm$ 8	49 [41-56]	0.88	41	0
POW1	252	49 $\pm$ 8	49 [41-54]	0.88	46	0
POW2	250	47 $\pm$ 8	41 [41-52]	0.92	58	0
POW3	244	47 $\pm$ 8	41 [41-54]	0.90	56	0
POW4	257	46 $\pm$ 8	41 [41-52]	0.90	63	0
<i>Fatigue (range 33.7 – 75.8)</i>						
Preoperatively	276	50 $\pm$ 10	49 [43-57]	0.94	16	2
POW1	252	53 $\pm$ 10	53 [49-59]	0.94	8	3
POW2	250	50 $\pm$ 9	49 [46-55]	0.93	10	1
POW3	244	49 $\pm$ 10	49 [43-55]	0.95	17	3
POW4	257	47 $\pm$ 10	49 [40-51]	0.96	24	1
<i>Sleep disorders (range 32 – 73.3)</i>						
Preoperatively	275	50 $\pm$ 8	50 [46-56]	0.84	4	0
POW1	252	50 $\pm$ 8	50 [46-54]	0.88	4	3
POW2	250	48 $\pm$ 8	48 [44-52]	0.86	4	1
POW3	242	48 $\pm$ 8	48 [41-52]	0.89	9	2
POW4	256	47 $\pm$ 9	48 [41-52]	0.87	12	2
<i>Ability to participate in social roles and activities (range 27.5 – 64.2)</i>						
Preoperatively	275	51 $\pm$ 11	52 [44-64]	0.95	6	29
POW1	252	44 $\pm$ 10	44 [37-52]	0.94	10	9
POW2	250	48 $\pm$ 10	48 [42-54]	0.94	6	15
POW3	243	50 $\pm$ 10	50 [44-58]	0.95	5	23
POW4	257	52 $\pm$ 10	52 [44-64]	0.95	3	30
<i>Pain interference (range 41.6 – 75.6)</i>						
Preoperatively	276	51 $\pm$ 11	50 [42-60]	0.98	47	5
POW1	252	58 $\pm$ 9	58 [54-64]	0.96	15	7
POW2	250	53 $\pm$ 9	56 [42-60]	0.96	28	2
POW3	243	51 $\pm$ 9	54 [42-56]	0.94	38	0
POW4	257	50 $\pm$ 9	50 [42-56]	0.97	47	3
<i>Pain intensity (range 0-10)</i>						
Preoperatively	276	2.6 $\pm$ 2.8	2 [0-4]	-	34	1
POW1	252	3.1 $\pm$ 2.1	3 [1-4]	-	10	0
POW2	250	2.2 $\pm$ 1.9	2 [1-3]	-	18	0
POW3	244	1.9 $\pm$ 1.9	1 [0-3]	-	27	1

POW4	257	1.6 ± 2.1	1 [0-2]	-	37	1
------	-----	-----------	---------	---	----	---

The floor effect represent the frequency of patients achieving the lowest PROMIS-29 t-score.

The ceiling effect represent the frequency of patients achieving the highest PROMIS-29 t-score.

A higher and a lower T-value indicates a superior health-reported quality of life in positive and negative domains, respectively.

Positive domains= Physical Function, Ability to participate in social roles and activities; *Negative domains*= Anxiety, Depression, Fatigue, Sleep disturbance, pain interference.

POW, postoperative week.

Supplemental Digital Content 4. Construct validity and responsiveness of PROMIS-29 physical and mental health summary scores – *Sensitivity analysis considering a minimally important difference (MID) of standard error of measure (SEM) of each baseline t-score.*

Hypotheses for construct validity	Mean difference (95% CI)			
	POW1	POW2	POW3	POW4
<i>Physical health summary score (MID 3.3)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.1 (0.9 to 3.1)	2.6 (0.7 to 4.6)	3.4 (1.6 to 5.5)	3.3 (1.4 to 5.3)
2. No 30day-complications vs. 30day-complications	2.3 (0.4 to 4.3)	2.4 (0.5 to 4.3)	2.7 (0.6 to 4.7)	2.9 (1.0 to 4.9)
3. Minimally invasive vs. open	1.8 (-0.4 to 4.1)	3.2 (1.1 to 5.4)	2.9 (0.6 to 5.2)	2.3 (0.1 to 4.5)
4. No new stoma vs. new stoma	1.4 (-0.9 to 3.7)	1.8 (0.4 to 4.0)	2.2 (0.2 to 4.6)	1.6 (0.6 to 3.9)
- No severe complications vs. severe complications	0.9 (-2.8 to 4.6)	1.7 (-1.9 to 5.5)	2.9 (1.0 to 6.8)	3.3 (0.3 to 6.9)
<i>Mental health summary score (MID 3)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	0.2 (-1.8 to 2.1)	2.3 (0.4 to 4.1)	2.9 (1.0 to 4.9)	2.2 (0.3 to 4.2)
2. No 30-day complications vs. 30-day complications	1.7 (-0.3 to 3.6)	2.4 (0.6 to 4.2)	2.6 (0.6 to 4.6)	2.1 (0.2 to 4.0)
3. Minimally invasive vs. Open	0.3 (-2.6 to 1.9)	1.2 (-0.9 to 3.2)	1.3 (-0.9 to 3.5)	1.1 (-1.1 to 3.3)
4. No new stoma vs. new stoma	1.7 (-0.6 to 4.0)	2.5 (0.4 to 4.6)	3.3 (1.0 to 5.5)	2.1 (0.2 to 4.3)
- No severe complications vs. severe complications	1.4 (-2.4 to 5.1)	2.3 (-1.0 to 5.7)	3.6 (-0.2 to 7.4)	3.0 (-0.6 to 6.6)
Hypotheses for responsiveness	Mean difference (95% CI)			
	Physical health summary score (MID 3.3)		Mental health summary score (MID 3)	
1. Baseline vs. POW1	9.9 (8.5 to 11.4)		3.1 (1.6 to 4.5)	
2. POW2 vs. POW1	3.9 (2.5 to 5.2)		3.1 (1.8 to 4.5)	
3. POW3 vs. POW2	2.7 (1.3 to 4.2)		1.6 (0.2 to 3.0)	
4. POW4 vs. POW3	2.4 (0.9 to 3.9)		1.8 (0.3 to 3.2)	

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID) of standard error of measure (SEM) of each baseline t-score.

LOS, length of stay; POW, postoperative week. Severe complications are defined as Clavien-Dindo ≥ 3.

Supplemental Digital Content 5. Construct validity and responsiveness of PROMIS-29 physical and mental health summary scores – *Sensitivity analysis considering a minimally important difference (MID) of half the standard deviation of each baseline t-score.*

Hypothesis for construct validity	Mean difference (95% CI)			
	POW1	POW2	POW3	POW4
<i>Physical health summary score (MID 5)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.1 (0.9 to 3.1)	2.6 (0.7 to 4.6)	3.4 (1.6 to 5.5)	3.3 (1.4 to 5.3)
2. No 30day-complications vs. 30day-complications	2.3 (0.4 to 4.3)	2.4 (0.5 to 4.3)	2.7 (0.6 to 4.7)	2.9 (1.0 to 4.9)
3. Minimally invasive vs. open	1.8 (-0.4 to 4.1)	3.2 (1.1 to 5.4)	2.9 (0.6 to 5.2)	2.3 (0.1 to 4.5)
4. No new stoma vs. new stoma	1.4 (-0.9 to 3.7)	1.8 (0.4 to 4.0)	2.2 (0.2 to 4.6)	1.6 (0.6 to 3.9)
- No severe complications vs. severe complications	0.9 (-2.8 to 4.6)	1.7 (-1.9 to 5.5)	2.9 (1.0 to 6.8)	3.3 (0.3 to 6.9)
<i>Mental health summary score (MID 4.5)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	0.2 (-1.8 to 2.1)	2.3 (0.4 to 4.1)	2.9 (1.0 to 4.9)	2.2 (0.3 to 4.2)
2. No 30day-complications vs. 30day complications	1.7 (-0.3 to 3.6)	2.4 (0.6 to 4.2)	2.6 (0.6 to 4.6)	2.1 (0.2 to 4.0)
3. Minimally invasive vs. open	0.3 (-2.6 to 1.9)	1.2 (-0.9 to 3.2)	1.3 (-0.9 to 3.5)	1.1 (-1.1 to 3.3)
4. No new stoma vs. new stoma	1.7 (-0.6 to 4.0)	2.5 (0.4 to 4.6)	3.3 (1.0 to 5.5)	2.1 (0.2 to 4.3)
- No severe complications vs. severe complications	1.4 (-2.4 to 5.1)	2.3 (-1.0 to 5.7)	3.6 (-0.2 to 7.4)	3.0 (-0.6 to 6.6)
Hypotheses for responsiveness	Mean difference (95% CI)			
	<i>Physical health summary score (MID 5)</i>		<i>Mental health summary score (MID 4.5)</i>	
1. Baseline vs. POW1	9.9 (8.5 to 11.4)		3.1 (1.6 to 4.5)	
2. POW2 vs. POW1	3.9 (2.5 to 5.2)		3.1 (1.8 to 4.5)	
3. POW3 vs. POW2	2.7 (1.3 to 4.2)		1.6 (0.2 to 3.0)	
4. POW4 vs. POW3	2.4 (0.9 to 3.9)		1.8 (0.3 to 3.2)	

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above the minimally important difference (MID).

LOS, length of stay; POW, postoperative week. Severe complications are defined as Clavien-Dindo ≥ 3.

Supplemental Digital Content 6. Construct validity of PROMIS-29 domains

Domains	POW1	POW2	POW3	POW4
<i>Physical Function</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	1.6 (0.2 to 3.5)	2.8 (1.0 to 4.7)	3.5 (1.6 to 5.5)	3.6 (1.7 to 5.5)
2. No 30day-complications vs. 30day-complications	2.3 (0.5 to 4.2)	2.6 (0.7 to 4.4)	2.7 (0.7 to 4.6)	3.1 (1.3 to 5.0)
3. Minimally invasive vs. open	2.2 (0.0 to 4.3)	3.2 (1.1 to 5.4)	2.9 (0.7 to 5.1)	2.4 (0.2 to 4.5)
4. No new stoma vs. New stoma	1.4 (-0.8 to 3.6)	1.9 (-0.3 to 4.1)	2.0 (0.2 to 4.3)	1.8 (-0.4 to 4.0)
- No severe complications vs. severe complications	1.3 (-2.4 to 4.9)	1.9 (-1.6 to 5.4)	2.5 (-1.1 to 6.1)	3.4 (-0.1 to 6.9)
<i>Anxiety</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-1.2 (-3.4 to 0.9)	-1.8 (-3.8 to 0.3)	-2.9 (-4.9 to -0.8)	-2.2 (-4.3 to -0.2)
2. No 30day-complications vs. 30day-complications	-1.1 (-3.2 to 1.1)	-2.5 (-4.5 to -0.4)	-2.4 (-4.5 to -0.4)	-2.4 (-4.4 to -0.4)
3. Minimally invasive vs. Open	-0.2 (-2.7 to 2.2)	0.8 (-1.6 to 3.1)	-0.8 (-3.1 to 1.6)	-0.1 (-2.4 to 2.2)
4. No new stoma vs. new stoma	-1.8 (-4.4 to 0.8)	-2.4 (-4.8 to 0.0)	-2.9 (-5.4 to -0.5)	-2.4 (-4.9 to 0.0)
- No severe complications vs. severe complications	-0.5 (-4.5 to 3.5)	-1.4 (-5.3 to 2.5)	-2.0 (-5.9 to 2.0)	-3.3 (-7.1 to 0.5)
<i>Depression</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-2.1 (-4.1 to -0.1)	-2.1 (-4.0 to -0.2)	-2.2 (-4.0 to -0.3)	-2.8 (-4.6 to -1.0)
2. No 30day-complications vs. 30day-complications	-1.0 (-3.0 to 0.9)	-1.9 (-3.8 to -0.0)	-2.0 (-3.9 to -0.2)	-1.9 (-3.7 to -0.1)
3. Minimally invasive vs. open	0.4 (-1.9 to 2.6)	0.3 (-1.9 to 2.5)	-0.2 (-2.4 to 1.9)	-0.5 (-2.6 to 1.5)
4. No new stoma vs. New stoma	-2.0 (-0.3 to 4.3)	-2.4 (-4.7 to -0.2)	-3.0 (-5.2 to -0.7)	-2.2 (-4.3 to -0.1)
- No severe complications vs. severe complications	1.2 (-2.6 to 5.0)	-0.5 (-4.1 to 3.1)	-2.2 (-5.8 to 1.4)	-1.5 (-4.9 to 1.8)
<i>Fatigue</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-1.3 (-3.6 to 1.0)	-3.4 (-5.5 to -1.2)	-4.3 (-6.6 to -2.0)	-4.1 (-6.3 to -1.8)
2. No 30day-complications vs. 30day-complications	-3.3 (-5.6 to -1.0)	-3.1 (-5.2 to -0.9)	-4.0 (-6.3 to -1.6)	-3.8 (-6.1 to -1.6)
3. Minimally invasive vs. open	0.1 (-2.5 to 2.8)	-2.0 (-4.5 to 0.4)	-2.2 (-4.9 to 0.5)	-2.5 (-5.1 to 0.1)
4. No new stoma vs. New stoma	-1.3 (-4.1 to 1.4)	-2.7 (-5.2 to -0.1)	-2.6 (-5.3 to 0.2)	-2.2 (-4.9 to 0.5)
- No severe complications vs. severe complications	-3.2 (-7.6 to 1.1)	-2.9 (-6.9 to 1.1)	-4.5 (-8.9 to -0.1)	-4.7 (-9.0 to -0.5)
<i>Sleep disturbance</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	1.5 (-0.5 to 3.5)	-0.2 (-2.1 to 1.7)	-0.6 (-2.5 to 1.3)	0.2 (-1.8 to 2.2)

2.	No 30day-complications vs. 30day-complications	-0.1 (-2.1 to 1.8)	-1.0 (-2.9 to 0.8)	-1.7 (-3.5 to 0.2)	-1.0 (-3.0 to 0.9)
3.	Minimally invasive vs. open	-0.4 (-2.7 to 1.9)	-0.2 (-2.3 to 2.0)	-0.3 (-2.5 to 1.9)	-0.2 (-2.4 to 2.1)
4.	No new stoma vs. New stoma	1.0 (-3.4 to 1.4)	-0.8 (-2.9 to 1.4)	-2.6 (-4.9 to -0.3)	-1.8 (-4.2 to 0.5)
-	No severe complications vs. severe complications	-1.4 (-5.4 to 2.5)	-1.5 (-5.0 to 2.1)	-3.7 (-7.3 to -0.2)	-1.1 (-4.8 to 2.6)
<i>Ability to participate in social roles and activities</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-0.1 (-2.7 to 2.5)	2.3 (-0.0 to 4.6)	2.6 (0.1 to 5.0)	1.1 (-1.3 to 3.5)
2.	No 30day-complications vs. 30day-complications	2.7 (0.2 to 5.3)	3.0 (0.7 to 5.3)	2.6 (0.2 to 5.1)	1.1 (-1.2 to 3.5)
3.	Minimally invasive vs. open	0.3 (-2.6 to 3.2)	2.5 (-0.1 to 5.2)	1.9 (-1.0 to 4.7)	0.7 (-2.0 to 3.3)
4.	No new stoma vs. New stoma	1.8 (-1.2 to 4.8)	2.1 (-0.6 to 4.8)	3.6 (0.7 to 6.5)	1.3 (-1.5 to 4.1)
-	No severe complications vs. severe complications	2.0 (-2.8 to 6.8)	3.7 (-0.7 to 8.0)	5.5 (0.1 to 10.0)	3.2 (1.1 to 7.5)
<i>Pain interference</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	3.1 (0.7 to 5.5)	-0.1 (-2.3 to 2.2)	-0.6 (-2.8 to 1.7)	-0.9 (-3.1 to 1.3)
2.	No 30day-complications vs. 30day-complications	0.1 (-2.3 to 2.5)	-0.7 (-2.9 to 1.5)	-1.1 (-3.3 to 1.1)	-1.4 (-3.6 to 0.8)
3.	Minimally invasive vs. open	1.2 (-1.5 to 3.9)	1.4 (1.2 to 4.0)	-1.1 (-3.7 to 1.5)	-1.3 (-3.8 to 1.2)
4.	No new stoma vs. New stoma	-1.11 (-3.9 to 1.7)	-1.9 (-4.5 to 0.7)	-2.6 (-5.2 to -0.0)	-1.4 (-3.9 to 1.2)
-	No severe complications vs. severe complications	1.2 (-3.5 to 5.9)	0.4 (-3.7 to 4.5)	-1.4 (-5.7 to 2.8)	1.0 (-5.0 to 3.1)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID) of 3 t-score points.

POW, postoperative week; LOS, length of stay. Severe complications are defined as Clavien-Dindo ≥ 3.

Supplemental Digital Content 7. Construct validity of PROMIS-29 domains – *Sensitivity analysis considering a minimally important difference (MID) of standard error of measure (SEM) of each baseline t-score.*

Domains	POW1	POW2	POW3	POW4
<i>Physical Function (MID 3)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.6 (0.2 to 3.5)	2.8 (1.0 to 4.7)	3.5 (1.6 to 5.5)	3.6 (1.7 to 5.5)
2. No 30day-complications vs. 30day-complications	2.3 (0.5 to 4.2)	2.6 (0.7 to 4.4)	2.7 (0.7 to 4.6)	3.1 (1.3 to 5.0)
3. Minimally invasive vs. open	2.2 (0.0 to 4.3)	3.2 (1.1 to 5.4)	2.9 (0.7 to 5.1)	2.4 (0.2 to 4.5)
4. No new stoma vs. New stoma	1.4 (-0.8 to 3.6)	1.9 (-0.3 to 4.1)	2.0 (0.2 to 4.3)	1.8 (-0.4 to 4.0)
- No severe complications vs. severe complications	1.3 (-2.4 to 4.9)	1.9 (-1.6 to 5.4)	2.5 (-1.1 to 6.1)	3.4 (-0.1 to 6.9)
<i>Anxiety (MID 2.8)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-1.2 (-3.4 to 0.9)	-1.8 (-3.8 to 0.3)	-2.9 (-4.9 to -0.8)	-2.2 (-4.3 to -0.2)
2. No 30day-complications vs. 30day-complications	-1.1 (-3.2 to 1.1)	-2.5 (-4.5 to -0.4)	-2.4 (-4.5 to -0.4)	-2.4 (-4.4 to -0.4)
3. Minimally invasive vs. Open	-0.2 (-2.7 to 2.2)	0.8 (-1.6 to 3.1)	-0.8 (-3.1 to 1.6)	-0.1 (-2.4 to 2.2)
4. No new stoma vs. new stoma	-1.8 (-4.4 to 0.8)	-2.4 (-4.8 to 0.0)	-2.9 (-5.4 to -0.5)	-2.4 (-4.9 to 0.0)
- No severe complications vs. severe complications	-0.5 (-4.5 to 3.5)	-1.4 (-5.3 to 2.5)	-2.0 (-5.9 to 2.0)	-3.3 (-7.1 to 0.5)
<i>Depression (MID 2.8)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-2.1 (-4.1 to -0.1)	-2.1 (-4.0 to -0.2)	-2.2 (-4.0 to -0.3)	-2.8 (-4.6 to -1.0)
2. No 30day-complications vs. 30day-complications	-1.0 (-3.0 to 0.9)	-1.9 (-3.8 to -0.0)	-2.0 (-3.9 to -0.2)	-1.9 (-3.7 to -0.1)
3. Minimally invasive vs. open	0.4 (-1.9 to 2.6)	0.3 (-1.9 to 2.5)	-0.2 (-2.4 to 1.9)	-0.5 (-2.6 to 1.5)
4. No new stoma vs. New stoma	-2.0 (-0.3 to 4.3)	-2.4 (-4.7 to -0.2)	-3.0 (-5.2 to -0.7)	-2.2 (-4.3 to -0.1)
- No severe complications vs. severe complications	1.2 (-2.6 to 5.0)	-0.5 (-4.1 to 3.1)	-2.2 (-5.8 to 1.4)	-1.5 (-4.9 to 1.8)
<i>Fatigue (MID 2.4)</i>				
1. Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-1.3 (-3.6 to 1.0)	-3.4 (-5.5 to -1.2)	-4.3 (-6.6 to -2.0)	-4.1 (-6.3 to -1.8)
2. No 30day-complications vs. 30day-complications	-3.3 (-5.6 to -1.0)	-3.1 (-5.2 to -0.9)	-4.0 (-6.3 to -1.6)	-3.8 (-6.1 to -1.6)
3. Minimally invasive vs. open	0.1 (-2.5 to 2.8)	-2.0 (-4.5 to 0.4)	-2.2 (-4.9 to 0.5)	-2.5 (-5.1 to 0.1)
4. No new stoma vs. New stoma	-1.3 (-4.1 to 1.4)	-2.7 (-5.2 to -0.1)	-2.6 (-5.3 to 0.2)	-2.2 (-4.9 to 0.5)
- No severe complications vs. severe complications	-3.2 (-7.6 to 1.1)	-2.9 (-6.9 to 1.1)	-4.5 (-8.9 to -0.1)	-4.7 (-9.0 to -0.5)

<i>Sleep disturbance (MID 3.2)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.5 (-0.5 to 3.5)	-0.2 (-2.1 to 1.7)	-0.6 (-2.5 to 1.3)	0.2 (-1.8 to 2.2)
2.	No 30day-complications vs. 30day-complications	-0.1 (-2.1 to 1.8)	-1.0 (-2.9 to 0.8)	-1.7 (-3.5 to 0.2)	-1.0 (-3.0 to 0.9)
3.	Minimally invasive vs. open	-0.4 (-2.7 to 1.9)	-0.2 (-2.3 to 2.0)	-0.3 (-2.5 to 1.9)	-0.2 (-2.4 to 2.1)
4.	No new stoma vs. New stoma	1.0 (-3.4 to 1.4)	-0.8 (-2.9 to 1.4)	-2.6 (-4.9 to -0.3)	-1.8 (-4.2 to 0.5)
-	No severe complications vs. severe complications	-1.4 (-5.4 to 2.5)	-1.5 (-5.0 to 2.1)	-3.7 (-7.3 to -0.2)	-1.1 (-4.8 to 2.6)
<i>Ability to participate in social roles and activities (MID 2.5)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-0.1 (-2.7 to 2.5)	2.3 (-0.0 to 4.6)	2.6 (0.1 to 5.0)	1.1 (-1.3 to 3.5)
2.	No 30day-complications vs. 30day-complications	2.7 (0.2 to 5.3)	3.0 (0.7 to 5.3)	2.6 (0.2 to 5.1)	1.1 (-1.2 to 3.5)
3.	Minimally invasive vs. open	0.3 (-2.6 to 3.2)	2.5 (-0.1 to 5.2)	1.9 (-1.0 to 4.7)	0.7 (-2.0 to 3.3)
4.	No new stoma vs. New stoma	1.8 (-1.2 to 4.8)	2.1 (-0.6 to 4.8)	3.6 (0.7 to 6.5)	1.3 (-1.5 to 4.1)
-	No severe complications vs. severe complications	2.0 (-2.8 to 6.8)	3.7 (-0.7 to 8.0)	5.5 (0.1 to 10.0)	3.2 (1.1 to 7.5)
<i>Pain interference (MID 1.6)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	3.1 (0.7 to 5.5)	-0.1 (-2.3 to 2.2)	-0.6 (-2.8 to 1.7)	-0.9 (-3.1 to 1.3)
2.	No 30day-complications vs. 30day-complications	0.1 (-2.3 to 2.5)	-0.7 (-2.9 to 1.5)	-1.1 (-3.3 to 1.1)	-1.4 (-3.6 to 0.8)
3.	Minimally invasive vs. open	1.2 (-1.5 to 3.9)	1.4 (1.2 to 4.0)	-1.1 (-3.7 to 1.5)	-1.3 (-3.8 to 1.2)
4.	No new stoma vs. New stoma	-1.1 (-3.9 to 1.7)	-1.9 (-4.5 to 0.7)	-2.6 (-5.2 to -0.0)	-1.4 (-3.9 to 1.2)
-	No severe complications vs. severe complications	1.2 (-3.5 to 5.9)	0.4 (-3.7 to 4.5)	-1.4 (-5.7 to 2.8)	1.0 (-5.0 to 3.1)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID) of standard error of measure (SEM) of each baseline t-score. *POW*, postoperative week; *LOS*, length of stay. Severe complications are defined as Clavien-Dindo ≥ 3.

Supplemental Digital Content 8. Construct validity of PROMIS-29 domains – *Sensitivity analysis considering a minimally important difference (MID) of half the standard deviation of each baseline t-score.*

Domains	POW1	POW2	POW3	POW4
<i>Physical Function (MID 4.5)</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	1.6 (0.2 to 3.5)	2.8 (1.0 to 4.7)	3.5 (1.6 to 5.5)	3.6 (1.7 to 5.5)
2. No 30day-complications vs. 30day-complications	2.3 (0.5 to 4.2)	2.6 (0.7 to 4.4)	2.7 (0.7 to 4.6)	3.1 (1.3 to 5.0)
3. Minimally invasive vs. open	2.2 (0.0 to 4.3)	3.2 (1.1 to 5.4)	2.9 (0.7 to 5.1)	2.4 (0.2 to 4.5)
4. No new stoma vs. New stoma	1.4 (-0.8 to 3.6)	1.9 (-0.3 to 4.1)	2.0 (0.2 to 4.3)	1.8 (-0.4 to 4.0)
- No severe complications vs. severe complications	1.3 (-2.4 to 4.9)	1.9 (-1.6 to 5.4)	2.5 (-1.1 to 6.1)	3.4 (-0.1 to 6.9)
<i>Anxiety (MID 4.5)</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-1.2 (-3.4 to 0.9)	-1.8 (-3.8 to 0.3)	-2.9 (-4.9 to -0.8)	-2.2 (-4.3 to -0.2)
2. No 30day-complications vs. 30day-complications	-1.1 (-3.2 to 1.1)	-2.5 (-4.5 to -0.4)	-2.4 (-4.5 to -0.4)	-2.4 (-4.4 to -0.4)
3. Minimally invasive vs. Open	-0.2 (-2.7 to 2.2)	0.8 (-1.6 to 3.1)	-0.8 (-3.1 to 1.6)	-0.1 (-2.4 to 2.2)
4. No new stoma vs. new stoma	-1.8 (-4.4 to 0.8)	-2.4 (-4.8 to 0.0)	-2.9 (-5.4 to -0.5)	-2.4 (-4.9 to 0.0)
- No severe complications vs. severe complications	-0.5 (-4.5 to 3.5)	-1.4 (-5.3 to 2.5)	-2.0 (-5.9 to 2.0)	-3.3 (-7.1 to 0.5)
<i>Depression (MID 4)</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-2.1 (-4.1 to -0.1)	-2.1 (-4.0 to -0.2)	-2.2 (-4.0 to -0.3)	-2.8 (-4.6 to -1.0)
2. No 30day-complications vs. 30day-complications	-1.0 (-3.0 to 0.9)	-1.9 (-3.8 to -0.0)	-2.0 (-3.9 to -0.2)	-1.9 (-3.7 to -0.1)
3. Minimally invasive vs. open	0.4 (-1.9 to 2.6)	0.3 (-1.9 to 2.5)	-0.2 (-2.4 to 1.9)	-0.5 (-2.6 to 1.5)
4. No new stoma vs. New stoma	-2.0 (-0.3 to 4.3)	-2.4 (-4.7 to -0.2)	-3.0 (-5.2 to -0.7)	-2.2 (-4.3 to -0.1)
- No severe complications vs. severe complications	1.2 (-2.6 to 5.0)	-0.5 (-4.1 to 3.1)	-2.2 (-5.8 to 1.4)	-1.5 (-4.9 to 1.8)
<i>Fatigue (MID 5)</i>				
1. Shorter LOS (LOS ≤ 4 days) vs. longer LOS (LOS > 4 days)	-1.3 (-3.6 to 1.0)	-3.4 (-5.5 to -1.2)	-4.3 (-6.6 to -2.0)	-4.1 (-6.3 to -1.8)
2. No 30day-complications vs. 30day-complications	-3.3 (-5.6 to -1.0)	-3.1 (-5.2 to -0.9)	-4.0 (-6.3 to -1.6)	-3.8 (-6.1 to -1.6)
3. Minimally invasive vs. open	0.1 (-2.5 to 2.8)	-2.0 (-4.5 to 0.4)	-2.2 (-4.9 to 0.5)	-2.5 (-5.1 to 0.1)
4. No new stoma vs.	-1.3 (-4.1 to 1.4)	-2.7 (-5.2 to -0.1)	-2.6 (-5.3 to 0.2)	-2.2 (-4.9 to 0.5)

New stoma					
-	No severe complications vs. severe complications	-3.2 (-7.6 to 1.1)	-2.9 (-6.9 to 1.1)	-4.5 (-8.9 to -0.1)	-4.7 (-9.0 to -0.5)
<i>Sleep disturbance (MID 4)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	1.5 (-0.5 to 3.5)	-0.2 (-2.1 to 1.7)	-0.6 (-2.5 to 1.3)	0.2 (-1.8 to 2.2)
2.	No 30day-complications vs. 30day-complications	-0.1 (-2.1 to 1.8)	-1.0 (-2.9 to 0.8)	-1.7 (-3.5 to 0.2)	-1.0 (-3.0 to 0.9)
3.	Minimally invasive vs. open	-0.4 (-2.7 to 1.9)	-0.2 (-2.3 to 2.0)	-0.3 (-2.5 to 1.9)	-0.2 (-2.4 to 2.1)
4.	No new stoma vs. New stoma	1.0 (-3.4 to 1.4)	-0.8 (-2.9 to 1.4)	-2.6 (-4.9 to -0.3)	-1.8 (-4.2 to 0.5)
-	No severe complications vs. severe complications	-1.4 (-5.4 to 2.5)	-1.5 (-5.0 to 2.1)	-3.7 (-7.3 to -0.2)	-1.1 (-4.8 to 2.6)
<i>Ability to participate in social roles and activities (MID 5.5)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	-0.1 (-2.7 to 2.5)	2.3 (-0.0 to 4.6)	2.6 (0.1 to 5.0)	1.1 (-1.3 to 3.5)
2.	No 30day-complications vs. 30day-complications	2.7 (0.2 to 5.3)	3.0 (0.7 to 5.3)	2.6 (0.2 to 5.1)	1.1 (-1.2 to 3.5)
3.	Minimally invasive vs. open	0.3 (-2.6 to 3.2)	2.5 (-0.1 to 5.2)	1.9 (-1.0 to 4.7)	0.7 (-2.0 to 3.3)
4.	No new stoma vs. New stoma	1.8 (-1.2 to 4.8)	2.1 (-0.6 to 4.8)	3.6 (0.7 to 6.5)	1.3 (-1.5 to 4.1)
-	No severe complications vs. severe complications	2.0 (-2.8 to 6.8)	3.7 (-0.7 to 8.0)	5.5 (0.1 to 10.0)	3.2 (1.1 to 7.5)
<i>Pain interference (MID 5.5)</i>					
1.	Shorter LOS (LOS ≤4 days) vs. longer LOS (LOS >4 days)	3.1 (0.7 to 5.5)	-0.1 (-2.3 to 2.2)	-0.6 (-2.8 to 1.7)	-0.9 (-3.1 to 1.3)
2.	No 30day-complications vs. 30day-complications	0.1 (-2.3 to 2.5)	-0.7 (-2.9 to 1.5)	-1.1 (-3.3 to 1.1)	-1.4 (-3.6 to 0.8)
3.	Minimally invasive vs. open	1.2 (-1.5 to 3.9)	1.4 (1.2 to 4.0)	-1.1 (-3.7 to 1.5)	-1.3 (-3.8 to 1.2)
4.	No new stoma vs. New stoma	-1.11 (-3.9 to 1.7)	-1.9 (-4.5 to 0.7)	-2.6 (-5.2 to -0.0)	-1.4 (-3.9 to 1.2)
-	No severe complications vs. severe complications	1.2 (-3.5 to 5.9)	0.4 (-3.7 to 4.5)	-1.4 (-5.7 to 2.8)	1.0 (-5.0 to 3.1)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above the minimally important difference (MID). POW, postoperative week; LOS, length of stay. Severe complications are defined as Clavien-Dindo ≥ 3.

Supplemental Digital Content 9. Responsiveness hypotheses.

Responsiveness Hypotheses	Mean difference (95% CI)
<i>Physical function</i>	
1. Baseline vs. POW1	10.1 (8.7 to 11.5)
2. POW2 vs. POW1	3.5 (2.2 to 4.9)
3. POW3 vs. POW2	2.7 (1.3 to 4.1)
4. POW4 vs. POW3	2.3 (0.8 to 3.7)
<i>Anxiety</i>	
5. Baseline vs. POW1	3.4 (1.9 to 5.0)
6. POW2 vs. POW1	-2.6 (-4.0 to -1.1)
7. POW3 vs. POW2	-1.0 (-2.5 to 0.5)
8. POW4 vs. POW3	-0.8 (-2.3 to 0.7)
<i>Depression</i>	
1. Baseline vs. POW1	0.8 (-0.6 to 2.2)
2. POW2 vs. POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs. POW2	0.3 (-1.0 to 1.7)
4. POW4 vs. POW3	-1.1 (-2.5 to 0.2)
<i>Fatigue</i>	
1. Baseline vs. POW1	-3.7 (-5.4 to -2.0)
2. POW2 vs. POW1	-2.6 (-4.2 to -1.0)
3. POW3 vs. POW2	-1.9 (-3.6 to -0.3)
4. POW4 vs. POW3	-1.9 (-3.6 to -0.2)
<i>Sleep disturbance</i>	
1. Baseline vs. POW1	-0.2 (-1.6 to 1.2)
2. POW2 vs. POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs. POW2	-0.7 (-2.1 to 0.6)
4. POW4 vs. POW3	-0.8 (-2.2 to 0.7)
<i>Ability to participate in social roles and activities</i>	
1. Baseline vs. POW1	6.9 (5.1 to 8.7)
2. POW2 vs. POW1	3.7 (2.1 to 5.4)
3. POW3 vs. POW2	2.2 (0.5 to 3.8)
4. POW4 vs. POW3	2.0 (0.3 to 3.7)
<i>Pain interference</i>	
1. Baseline vs. POW1	-7.0 (-8.7 to -5.3)
2. POW2 vs. POW1	-4.7 (-6.2 to -3.1)
3. POW3 vs. POW2	-2.1 (-3.6 to -0.6)
4. POW4 vs. POW3	-1.7 (-3.2 to -0.2)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID) of 3 t-score points.

POW, postoperative week.

Supplemental Digital Content 10. Responsiveness hypotheses – *Sensitivity analysis* considering a minimally important difference (MID) of standard error of measure (SEM) of each baseline t-score.

Responsiveness Hypotheses	Mean difference (95% CI)
<i>Physical function (MID 3)</i>	
1. Baseline vs. POW1	10.1 (8.7 to 11.5)
2. POW2 vs. POW1	3.5 (2.2 to 4.9)
3. POW3 vs. POW2	2.7 (1.3 to 4.1)
4. POW4 vs. POW3	2.3 (0.8 to 3.7)
<i>Anxiety (MID 2.8)</i>	
1. Baseline vs. POW1	3.4 (1.9 to 5.0)
2. POW2 vs. POW1	-2.6 (-4.0 to -1.1)
3. POW3 vs. POW2	-1.0 (-2.5 to 0.5)
4. POW4 vs. POW3	-0.8 (-2.3 to 0.7)
<i>Depression (MID 2.8)</i>	
1. Baseline vs. POW1	0.8 (-0.6 to 2.2)
2. POW2 vs. POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs. POW2	0.3 (-1.0 to 1.7)
4. POW4 vs. POW3	-1.1 (-2.5 to 0.2)
<i>Fatigue (MID 2.4)</i>	
1. Baseline vs. POW1	-3.7 (-5.4 to -2.0)
2. POW2 vs. POW1	-2.6 (-4.2 to -1.0)
3. POW3 vs. POW2	-1.9 (-3.6 to -0.3)
4. POW4 vs. POW3	-1.9 (-3.6 to -0.2)
<i>Sleep disturbance (MID 3.2)</i>	
1. Baseline vs. POW1	-0.2 (-1.6 to 1.2)
2. POW2 vs. POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs. POW2	-0.7 (-2.1 to 0.6)
4. POW4 vs. POW3	-0.8 (-2.2 to 0.7)
<i>Ability to participate in social roles and activities (MID 2.5)</i>	
1. Baseline vs. POW1	6.9 (5.1 to 8.7)
2. POW2 vs. POW1	3.7 (2.1 to 5.4)
3. POW3 vs. POW2	2.2 (0.5 to 3.8)
4. POW4 vs. POW3	2.0 (0.3 to 3.7)
<i>Pain interference (MID 1.6)</i>	
1. Baseline vs. POW1	-7.0 (-8.7 to -5.3)
2. POW2 vs. POW1	-4.7 (-6.2 to -3.1)
3. POW3 vs. POW2	-2.1 (-3.6 to -0.6)
4. POW4 vs. POW3	-1.7 (-3.2 to -0.2)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID) of 2.7 t-score points.

POW, postoperative week.

Supplemental Digital Content 11. Responsiveness hypotheses – *Sensitivity analysis* considering a minimally important difference (MID) of half the standard deviation of each baseline t-score.

Responsiveness Hypotheses	Mean difference (95% CI)
<i>Physical function (MID 4.5)</i>	
1. Baseline vs POW1	10.1 (8.7 to 11.5)
2. POW2 vs POW1	3.5 (2.2 to 4.9)
3. POW3 vs POW2	2.7 (1.3 to 4.1)
4. POW4 vs POW3	2.3 (0.8 to 3.7)
<i>Anxiety (MID 4.5)</i>	
1. Baseline vs POW1	3.4 (1.9 to 5.0)
2. POW2 vs POW1	-2.6 (-4.0 to -1.1)
3. POW3 vs POW2	-1.0 (-2.5 to 0.5)
4. POW4 vs POW3	-0.8 (-2.3 to 0.7)
<i>Depression (MID 4)</i>	
1. Baseline vs POW1	0.8 (-0.6 to 2.2)
2. POW2 vs POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs POW2	0.3 (-1.0 to 1.7)
4. POW4 vs POW3	-1.1 (-2.5 to 0.2)
<i>Fatigue (MID 5)</i>	
1. Baseline vs POW1	-3.7 (-5.4 to -2.0)
2. POW2 vs POW1	-2.6 (-4.2 to -1.0)
3. POW3 vs POW2	-1.9 (-3.6 to -0.3)
4. POW4 vs POW3	-1.9 (-3.6 to -0.2)
<i>Sleep disturbance (MID 4)</i>	
1. Baseline vs POW1	-0.2 (-1.6 to 1.2)
2. POW2 vs POW1	-1.6 (-3.0 to -0.2)
3. POW3 vs POW2	-0.7 (-2.1 to 0.6)
4. POW4 vs POW3	-0.8 (-2.2 to 0.7)
<i>Ability to participate in social roles and activities (MID 5.5)</i>	
1. Baseline vs POW1	6.9 (5.1 to 8.7)
2. POW2 vs POW1	3.7 (2.1 to 5.4)
3. POW3 vs POW2	2.2 (0.5 to 3.8)
4. POW4 vs POW3	2.0 (0.3 to 3.7)
<i>Pain interference (MID 5.5)</i>	
1. Baseline vs POW1	-7.0 (-8.7 to -5.3)
2. POW2 vs POW1	-4.7 (-6.2 to -3.1)
3. POW3 vs POW2	-2.1 (-3.6 to -0.6)
4. POW4 vs POW3	-1.7 (-3.2 to -0.2)

Data are the mean difference (95% CI) in the t-score.

Hypothesis supported by the data are bolded. Mean differences were considered relevant if above a minimally important difference (MID).

POW, postoperative week.