American Society for Enhanced Recovery (ASER) and Perioperative Quality Initiative (POQI) Joint Consensus Statement on Patient Reported Outcomes (PROs) within an Enhanced Recovery Pathway


**Supplemental Digital Content**

**Supplemental Table 1.** Clinical studies that focus on patient reported outcomes (PROs) within the context of an enhanced recovery program (ERP).

<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Journal</th>
<th>PROM</th>
<th>Surgical Type</th>
<th>Summary of results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delaney 2003</td>
<td>Dis Colon Rectum</td>
<td>Cleveland Clinic Global Quality of Life Questionaire</td>
<td>Laparotomy &amp; Intestinal Resection</td>
<td>There was no difference between enhanced pathway and traditional patients in quality of life after surgery.</td>
</tr>
<tr>
<td>Jones 2013</td>
<td>Br J Surg</td>
<td>EQ5D</td>
<td>Open Liver Resection</td>
<td>ERP vs standard care. Quality of life was significantly better in ERP group.</td>
</tr>
<tr>
<td>King 2006</td>
<td>Colorectal Dis</td>
<td>EORTC QLQ-C30 QLQ-CR38</td>
<td>Colorectal cancer</td>
<td>ERP vs historical controls: There was no significant difference in quality of life outcomes</td>
</tr>
<tr>
<td>Larsen 2008</td>
<td>Acta Orthop</td>
<td>EQ-5D</td>
<td>Knee/Hip Arthroplasty</td>
<td>RCT: control vs accelerated perioperative care. EQ-5D scores were higher in both groups at 3 months compared to baseline, but the intervention group had better quality of life scores.</td>
</tr>
<tr>
<td>Larsen 2010</td>
<td>Health Qual Life Outcomes</td>
<td>EQ-5D, SF36</td>
<td>Hip Arthroplasty</td>
<td>Fast track vs population norm data. Hip surgery patients reached population normal levels for HRQOL @ 3 months and exceeded population normal @ 12 months.</td>
</tr>
</tbody>
</table>
**Supplemental Table 1 (continued).** Clinical studies that focus on patient reported outcomes (PROs) within the context of an enhanced recovery program (ERP).

<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Journal</th>
<th>PROM</th>
<th>Surgical Type</th>
<th>Summary of results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larsen 2012</td>
<td>Knee Surg Sports Traumatol Arthosc</td>
<td>EQ 5D, SF 36</td>
<td>Knee Arthroplasty</td>
<td>Fast track vs. population norm data. Patients were below population norms at 4 and 12 months. No or mild pain and good functional ability @ 4 months were associated with high HRQOL at later follow up</td>
</tr>
<tr>
<td>Lemanu 2013</td>
<td>Br J surg</td>
<td>Surgical Recovery Scale</td>
<td>Lap. Sleeve gastrectomy</td>
<td>ERAS vs control. No difference in postoperative fatigue between groups.</td>
</tr>
<tr>
<td>Recart 2005</td>
<td>J Endourol</td>
<td>Quality of Recovery Score</td>
<td>Lap Nephrectomy</td>
<td>Conventional Vs fast track. Quality of recovery scores similar between groups</td>
</tr>
<tr>
<td>Shida 2015</td>
<td>BMC Cancer</td>
<td>QoR-40</td>
<td>Colorectal cancer</td>
<td>ERAS patients. POD#1 and #3 scores significantly decreased with return to baseline on POD#6.</td>
</tr>
<tr>
<td>Wang 2010</td>
<td>J Gastrointest Surg</td>
<td>Spitzer Index</td>
<td>Gastric Cancer</td>
<td>Fast track vs Conventional. Fast track surgery was associated with higher quality of life score on hospital discharge</td>
</tr>
</tbody>
</table>

QLQ-C30: European Organization for Research & Treatment of Cancer Quality of Life Questionnaire C30
QLQ-CR38: European Organization for Research & Treatment of Cancer Quality of Life Questionnaire Colorectal 38
QLQ-CR28: European Organization for Research & Treatment of Cancer Quality of Life Questionnaire – Colorectal 28
SF-36: Short Form 36
MDASI-GI: Gastrointestinal version of MD Anderson Symptom Inventory
MDASI: MD Anderson Symptom Inventory
RCT: Randomized controlled trial
HRQOL: Health related quality of life
ERAS: Enhanced recovery after surgery
Supplemental Text 1

Summary of PROMs of Potential Relevance to Enhanced Recovery Pathways

Quality of Recovery Score (QoR) - QoR-9, QoR-15, QoR-40

The quality of recovery scores, QoR-9, QoR-15, QoR-40 were reported by Myles and colleagues in several studies. The goal of the QoR score is to provide a valid, reliable, and responsive measure of quality of recovery after anesthesia and surgery. The most comprehensive measure, the QoR-40, covers five health dimensions related to mental and physical well-being: 1. patient support, 2. emotions, 3. comfort, 4. physical independence and 5. pain, each scored on a zero to ten point scale. The QoR-40 showed superior validity and reliability in comparison to the QoR-9, however it requires approximately ten minutes to administer. Alternatively, the QoR-15 can be completed in 3 minutes and evidence supports its validity, reliability, responsiveness and feasibility in surgical patients in clinical practice. All QoR scores ask patients to evaluate their health with a 24-hour period, which makes them an attractive instrument for the immediate perioperative period. After colorectal surgery, QoR-40 scores were found to drop significantly on postoperative day 1, with significant improvement by postoperative day 3 and return to baseline on postoperative day 6. Compared to a variety of other patient centered tools, the QoR scores have shorter recall periods (24-hours) allowing for their use in the dynamic immediate postoperative phase when most ERP interventions continue.
WHODAS 2.0

The World Health Organization – Disability Assessment Scale 2.0 (WHODAS 2.0) is directly linked to the structural concepts with the World Health Organization's (WHO) International Classification of Functioning, Disability and Health, more commonly known as ICF. 17 Disability is defined by the WHO as a difficulty in functioning at the body, person, or societal level. Disability occurs in one or more life domains, as experienced by an individual with health conditions in interaction with contextual factors. WHODAS 2.0 follows a biopsychosocial model of health and covers 6 domain functions: 1. Cognition: understanding and communication, 2. Mobility: moving and getting around, 3. Self-care: hygiene, dressing, eating and staying alone, 4. Getting along: interacting with other people, 5. Life activities: domestic responsibilities, leisure, work and school, and 6. Participation: joining in community activities. WHODAS 2.0 is a generic assessment instrument for health and disability and can be used across all diseases, including mental, neurological and addictive disorders. It is short, simple and easy to administer (5-20 minutes). It has application in both clinical and general population settings. It is a tool that can be used to produce standardized disability levels and profiles applicable across cultures in all adult populations.19

Recently, the WHODAS 2.0 has shown adequate validity, reliability and responsiveness in a diverse surgical population.18 Five-hundred patients were assessed using the WHODAS 2.0 instrument following surgery. The WHODAS 2.0 correlated with QoR scores at 30 days, and measures of pain interference and physical function at 3, 6, and 12 months after surgery. Patients with increased hospital length of stay or complications
within the first 30 days correlated with a new disability in a life domain. This initial validation within the perioperative setting further supports its possible use within an ERP.

PROMIS

In 2004, the National Institutes of Health began the development of a system of PROs in order to overcome barriers to large scale clinical and research use of patient centered outcomes. The Patient Reported Outcome Measurement Information System (PROMIS- www.healthmeasures.net) leverages modern psychometric principles in order to provide a precise and widely applicable system of PROs. PROMIS measures are administered as item banks that are grouped under the three domains: 1. physical, 2. mental, 3. and social health. Each item bank underwent rigorous development utilizing item response theory (IRT) that maximizes precision in each item bank, increases flexibility and allows for tailored administration. Additionally, PROMIS can utilize Computer Adaptive Testing (CAT) through which questions are selected based on a patient’s previous answer. CAT adds the benefit of minimizing the number of questions to be answered without sacrificing reliability in the scores produced. A critical benefit of PROMIS is its use of a standardized metric, the T score. This is normalized to the general population and allows providers to longitudinally “speak the same language” across a variety of care settings. In institutions without the capability for CAT, PROMIS measures also are available in short form item banks (e.g. depression, pain interference) or short form profile instruments (e.g. PROMIS-29).
PROMIS measures are being integrated rapidly by surgical services (e.g. orthopedics, oncological surgery) and represent a cutting edge opportunity for pain physicians to influence rational evidence-based pain care. A scoping review characterized 21 publications where PROMIS measures were used in the perioperative setting. The authors applauded the utility of PROMIS measures to provide standardized, accurate and efficiently captured patient constructs. A PROMIS profile instrument (PROMIS 29-a non-computer adaptive profile) was used in an interdisciplinary opioid reduction program in patients preparing to undergo spine surgery where significant benefits in pain interference occurred throughout the perioperative period.

Additionally, when compared to measures such as WHODAS and EQ5D, PROMIS has displayed similar performance in numerous populations. Numerous PROMIS item banks (e.g. Pain Interference, Depression, Sleep Disturbance, etc.) have also shown to be equivalent if not superior compared to reference legacy instruments (e.g. Brief Pain Inventory, Center for Epidemiological Studies Depression Scale, Pittsburgh Sleep Quality Index). PROMIS measures, indexed over a 7-day period, provide the opportunity to assess the impact of ERP interventions and also the ability to act upon biopsychosocial variables that affect recovery in the immediate and sub-acute post discharge phases. Such frequent assessment would allow for construction of expected recovery trajectories where early intervention may further enhance function. PROMIS allows for users to tailor domain measures based on what health status measures they wish to assess. (Table 6) This allows for a tailored approached in a condition specific manner unlike generic measures (e.g. SF-36, EQ5D). While a good prospect for any
PRO program, future work is needed to further establish the utility of PROMIS measures in surgical population with or without an ERP.

EQ-5D

The EuroQol 5 dimension questionnaire (EQ-5D) is one of the mostly commonly used generic questionnaires to measure health related quality of life.\(^{30}\) The questionnaire covers five domains: 1. mobility, 2. self-care, 3. usual activities, 4. pain/discomfort, and 5. anxiety/depression. A patient grades their level of disability on a three-grade scale: severe, moderate or none. The EQ-5D-5L asks the same questions but with a 5-point scale instead of a 3-point scale. Conceptually, the EQ-5D was created with a holistic view of health, which is comprised of medical, physical independence, emotional and social functioning components. The questionnaire includes both positive (well-being) and negative (illness) questions. The EQ-5D combines both a questionnaire and a visual analog scale – EQ-VAS. The EQ-5D asks patients to rate their health status "today."

The EQ-5D is used in the National Health System (NHS) in England for assessment of patient outcomes after specific surgical procedures.\(^{31}\) The NHS has been collecting EQ-5D information since 2009, and this represents an effort to measure patient-reported health in several ways. Between April and June 2016, an increase in general health was recorded for 49% of patients after groin hernia surgery and 47% of patients after varicose vein surgery as measured using the EQ-5D index.\(^{32}\) The National Joint Registry offers one model for the use of PROMs in comparative effectiveness research.
EQ-5D score was higher at 6 months for unicompartmental knee arthroplasty (UKA) compared to total knee arthroplasty (TKA). UKA patients (n=3519) were more likely to achieve excellent results and be highly satisfied compared to TKA patients (n=10557). These authors concluded that the high revision rate of UKA may not be because of poorer clinical outcome.  

Overall, the EQ-5D is a widely used instrument internationally to assess numerous quality indices. However, some studies suggest that EQ-5D has limited content validity, construct validity and responsiveness in the context of surgical recovery. This instrument is not very discriminative and has a significant ceiling effect when used after surgery, particularly with abdominal and thoracic surgery. Future work in the perioperative and ERP arena must also focus on its use to provide actionable outcomes to enhance recovery instead of its sole use at remote time point distant from surgical intervention.

**Short Form – 36 Health Survey**

The Short Form – 36 Health Survey (SF-36) was created in 1992 by the Medical Outcomes Study as part of the RAND Corporation. The SF-36 was designed for use in clinical practice, research, health policy evaluations and general population surveys. SF-36 was built on the premise that good medical care should result in a more "effective life" and preserve function and well being. SF-36 can be self-administered, collected via telephone or in-person interview. One potential critique of the SF-36 is that it was
designed for use in medical populations; however, there are studies contributing
evidence for the measurement properties of the SF-36 in surgical populations.\textsuperscript{38}

The SF-36 contains 8 health concepts: 1. Limitation of physical activities – health
problems, 2. Limitation of social activities - physiological/emotional, 3. Limitation – usual
Limitation – usual role activities/emotional problems, 7. Vitality (energy/fatigue) and 8.
General health perceptions. The SF-36 has been employed in thousands of studies,
undergone hundreds of separate validations, and translated into more than 50
languages/cultures. The SF-36 contains items that are queried over various timeframes
such as “compared to one year ago,” “on a typical day,” and “during the last 4 weeks.”
Several studies have made use of the SF-36 in surgical patients where it was commonly
administered at remote time points, such as 30 days postoperatively and beyond.\textsuperscript{39-41} In
one example, the SF-36 did not find differences in patients who had open versus
laparoscopic abdominal surgery.\textsuperscript{38,42} While this may relate to the validity of the SF-36 in
ERPs, it suggests that measurement during the immediate post discharge phase
throughout standard surgical follow up appointments (e.g. 30 days) is possibly needed
to detect possible opportunities for further enhancement of recovery.

Short Form 12 and Short Form 1 Health Survey

The Short Form 12 (SF-12) and Short Form (SF-1) are both derivatives of the SF-36
that have been used in clinical practice to assess health related quality of life. The SF-
12 questionnaire has been developed to calculate and reproduce the physical component score and mental component score of the SF-36. The SF-12 has been validated in the general population, and patients following myocardial infarction, stroke, and on dialysis. The SF-12 has been used for surgical patients undergoing spine, renal and major surgery.

The SF-1 utilized the first question of the SF-36 to assess patients HRQOL: “In general, would you say your health is: Excellent, Very Good, Good, Fair of Poor?” (Gill) The SF1 has been used as a general indicator of self-reported wellbeing, an indicator for future health care use and mortality, and is in the Australian Health Survey. Although the SF1 is straightforward in its utility, it does not evaluate the mental, physical or social domains of a patient’s quality of life.

**EORTC QLQ- C30**

The European Organization for Research and Treatment of Cancer, QoL C30 (EORTC QLQ-C30) consists of 9 multi item scales. The EORTC-QLQ-C30 is currently used in all major oncology trials in Europe as a measure of quality of life. The survey incorporate 5 functional scales (physical, role, cognitive, emotional, social), 3 symptom scales (fatigue, pain and nausea/vomiting), global heath and a quality of life scale. The average time to perform the EORTC QLQ-C30 is 11 minutes. A number of symptoms that are commonly reported by cancer patients, e.g. dyspnea, loss of appetite, insomnia, constipation, and diarrhea, are included in the survey. The survey also
inquires about the financial impact of the disease. The EORTC-QLQ-C30 was found to contain a high number of meaningful measures of recovery.\textsuperscript{34}

\textit{Other instruments}

There are many other instruments that can be used to measure patient reported outcomes after surgery. These include the Postoperative Recovery Index (PORI)\textsuperscript{54} and the Gastrointestinal Quality of Life Index (GIQLI)\textsuperscript{55} as examples. The PORI is a quality of recovery scoring system that is self-reported and multi-dimensional. This measure has applicability across various surgeries and surgical settings, from immediately post-surgery throughout discharge and covering the first 30 days of recovery. The GIQLI was an instrument designed to measure quality of life specifically for patients with gastrointestinal disease.
Supplemental References


