RESEARCH

Open Access



Joshua M. Bonsel^{1*†}, Ademola J. Itiola^{2†}, Anouk S. Huberts³, Gouke J. Bonsel⁴ and Hannah Penton⁵

Abstract

Background Patient-reported outcome measures (PROMs) provide invaluable information on patients' health outcomes and can be used to improve patient-related outcomes at the individual, organizational and policy levels. This systematic review aimed to a) identify contemporary applications and synthesize all evidence on the use of PROMs in these contexts and b) to determine characteristics of interventions associated with increased effectiveness.

Methods Five databases were searched for studies providing quantitative evidence of the impact of PROM interventions. Any study design was permitted. An overall benefit (worsening) in outcome was defined as a statistically significant improvement (deterioration) in either a PROM, patient-reported experience measure or clinical outcome. Study quality was assessed using the Effective Public Healthcare Panacea Project's Quality Assessment Tool for Quantitative Studies. A narrative synthesis was conducted.

Results Seventy-six studies of the 11,121 articles identified met the inclusion criteria. At the individual level, 10 (43%) of 23 studies that fed back PROMs to the patient or healthcare provider showed an improvement in outcome. This percentage increased in studies which used PROMs to monitor disease symptoms and linked these to care-pathways: 17 (68%) of 25 studies using this mechanism showed an improvement. Ten (71%) of 14 studies using PROMs to screen for disease found a benefit. The monitoring and screening approach was most effective using PROMs covering cancer-related, depression and gastro-intestinal symptoms. Three studies found that the mere collection of PROMs resulted in improved outcomes. Another three studies used PROMs in decision aids and found improved decision quality.

At the organizational/policy level, none of the 4 studies that used PROMs for benchmarking found a benefit. The three studies that used PROMs for in-depth performance analyses and 1 study in a plan-do-study-act (PDCA) cycle found an improvement in outcome.

Studies employing disease-specific PROMs tended to observe improved outcomes more often. There are concerns regarding the validity of findings, as studies varied from weak to moderate quality.

Conclusions The use of PROMs at the individual level has matured considerably. Monitoring/screening applications seem promising particularly for diseases for which treatment algorithms rely on the experienced symptom burden by patients. Organizational/policy-level application is in its infancy, and performance evaluation via in-depth analyses

[†]Joshua M. Bonsel and Ademola J. Itiola shared equal first authorship.

*Correspondence: Joshua M. Bonsel j.bonsel@erasmusmc.nl Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

and PDCA-cycles may be useful. The findings of this review may aid stakeholders in the development and implementation of PROM-interventions which truly impact patient outcomes.

Keywords Patient-Reported Outcome Measures, Quality of care, Patient outcomes, Feedback, Monitoring, Benchmarking, Routine outcome measurement

Background

Patient-reported outcomes measures (PROMs) are considered an invaluable tool to capture information on patients' health outcomes, including expectations and values. Two types of PROMs exist, namely generic and disease-specific PROMs [1]. Generic PROMs aim to measure a health outcome from an overarching perspective, allowing for comparison between different diseases and a general judgement on the severity. These measures are often multi-dimensional; examples include measures of overall Quality of Life (e.g., EQ-5D) or well-being (e.g., WHO-5) [2, 3]. Disease-specific PROMs aim to measure these concepts, the symptom burden and functional status associated with a disease or a group of diseases [4].

PROMs were introduced to complement clinical outcome measures in studies assessing the (cost-) effectiveness of new clinical interventions. However, their application has broadened, including the role as outcome indicator in clinical practice alongside traditional indicators such as mortality and prevalence/ incidence [5]. This movement is adopted by medical science and leading institutions like the Organisation for Economic Co-operation and Development, which conform to the principle that assessing health system performance starts by assessment of patient-related outcomes [6]. It is pragmatic to distinguish three levels of intended use: the individual (micro-), organizational (meso-) and policy (macro-) level [7].

At the micro-level, PROMs are used at the patientencounter level. Several systematic reviews revealed evidence that using PROMs at the micro-level has a modest beneficial impact on patient-related outcomes [8-15]. The key idea is that a patient fills out a PROM once or multiple times, and the results are fed back to the patient or clinician [15]. Greenhalgh et al. has outlined the underlying theory how PROMs may be useful at this level: the feedback of PROMs may alter the decision-making process, and initiate a change to clinical practice [16]. Several examples exist: firstly, the feedback of PROMs to patient and provider can aid in communicating symptoms which may otherwise remain unnoticed [17, 18]. Another example are novel digital patient-decision systems using PROMs, which develop rapidly parallel to digital technology (e.g., apps, e-portals, and dashboards) [19].

Aggregated PROMs can be used to inform the healthcare system at the organizational (meso-) and health system (macro-) level, respectively. Evidence of the impact of PROMs use at the meso-/macro-level is scarce, and a recent review did not find a clear impact on patient outcomes [8, 20]. The key idea at this level is that aggregated PROMs can guide the (continuous) improvement of healthcare provided by a group of clinicians, hospital or even country [21]. Their role in orthopedic surgery may illustrate their potential. At the meso-level, an orthopedic surgery unit in a hospital may use PROMs to improve local policy on eligibility criteria for surgical treatment, to rationalize pain killing strategies, or to compare performance across surgeons on a monthly basis [22]. At the macro-level, PROMs results according to hospital, region, nation, or otherwise may be presented in a standardized form (both in epidemiological and graphical meaning), inviting for a process of feedback, analysis of drivers, and if possible subsequent improvement [21]. This mechanism is often referred to as benchmarking and is thought to demonstrate performance differences among providers, facilitate more in-depth clinical audits, and inform decision-making, and is a potentially effective method to improve the quality of care [23, 24]. An example which aimed to encourage benchmarking is the NHSprogramme in the UK on certain surgical procedures. This program publicly published PROMs for varicose vein, groin hernia, and hip and knee arthroplasty surgery; as of 2017 PROMs are only collected for hip/knee surgery [25]. This program also aimed to incentivize patients to select the assumed best provider, however, available evidence does not support this pathway [21, 26].

We think a contemporary review is warranted because it remains unknown why certain PROMsinterventions are more effective than others [8, 11]. Certain mechanisms underpinning the interventions may contribute to increased effectiveness. For example, a critical step to transform a suboptimal PROM level, i.e. a patient value below a particular threshold, into an improved outcome may be to link this observation to a care pathway. The doctor may receive an alert inviting her/him to check the situation. This approach seems promising in disease areas where symptom monitoring along with treatment tailoring is common practice, e.g., gastroenterology, rheumatology, and oncology [27, 28]. In this systematic review, we aim to identify contemporary evidence of the impact of the use of PROMs at the micro-, meso- and macro-level on patient outcomes. Our second aim is to identify and describe characteristics of the intervention and PROMs used which may contribute to an increased chance for success.

Methods

The present systematic review was registered in PROSPERO under record 2022 CRD42022333400. This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (2020) when applicable [29].

Data sources and search strategy

The following databases were searched: MEDLINE, Embase, Web of Science Core Collection, Cochrane CENTRAL Register of trials, and Google Scholar from database inception to August 24, 2023 for studies that reported the use of PROMs to improve quality of care. The final search was developed and refined through an iterative process and consisted of 3 blocks, namely: (a) various terms for PROMs, (b) various terms for quality, effectiveness and outcomes, and (c) mechanisms through which PROMs may be used to benefit healthcare (e.g., feedback, monitoring, dashboards and plan-do-checkact (PDCA) cycles) (Supplementary Material 1). A PDCA-cycle is a commonly used framework to guide the continuous improvement of healthcare and services provided [30]. Additional studies were identified by screening the references of included articles.

Study selection

Studies were eligible that (a) provided evidence on the impact of an intervention, (b) using a previously validated PROM, (c) which reported at least one quantitative outcome per the definition described below. Any study design was permitted. Studies were excluded if (a) the full-text could not be retrieved and/or only a conference abstract was available; (b) the study was conducted as a pilot; (c) there was no comparator or pre-intervention comparison; (d) the PROM was used to select patients for another type of intervention; (e) the article was not available in English. Two reviewers (JB and AI) independently screened all titles and abstracts obtained from the search and applied the inclusion criteria to eligible studies. Any disagreements regarding the inclusion of studies was discussed between the two reviewers and were resolved by consensus.

Outcome definition

We defined the potential impact of a PROM-intervention on patient-related outcomes using the Donabedian framework [31]. To evaluate the quality of healthcare or impact of an intervention, contemporary guidelines place emphasis on outcome measures which reflect the impact on the health status of patients [32]. Typically, these outcomes are of quantitative nature and are collected at the patient-level. We discerned three types of outcomes measured based on previous reviews, namely (1) PROMs, (2) patient reported experiences measures (PREMs) and (3) clinical outcomes. Outcome measures were categorized according to the dimensions/items into overarching groups based on the identified studies, e.g., Health-Related Quality of Life (HRQoL), physical functioning, mental functioning, and symptom burden. Similarly, this was done for PREMs (e.g., satisfaction) and clinical outcomes (e.g., readmissions).

A study was judged to have found an overall benefit (or a detriment/harm) if any of the above-mentioned outcomes improved (worsened) up to statistical significance. As patient-related outcomes may be specific to the intended use and medical domain, we did not attribute weight to a specific type of outcome. Studies often contained multiple comparisons through analysis of dimensions or even items separately. This approach inflates testing, increasing the potential of a type I error. Therefore, we required at least 2 subdomain/single-items to reach statistical significance to qualify the impact as a benefit or detriment, unless outcomes were defined as primary outcome a priori.

In accordance with previous reviews, process of care measures (e.g., number of symptoms discussed) were extracted, but were considered to mediate outcomes described above [14].

Data extraction and quality assessment

The following data were extracted from eligible studies by one of the reviewers (JB or AI): authors, country, setting, study design, sample, PROMs used, description of intervention using PROMs, co-interventions, training offered on the intervention and/or interpretation of PROM, all primary and secondary outcome measures and their quantification.

Two reviewers (JB and AH) independently assessed the methodological quality of included studies using the Effective Public Healthcare Panacea Project's Quality Assessment Tool for Quantitative Studies [33]. The tool was considered the most appropriate for this systematic review as it covers various study designs and public health interventions. Domains assessed using the tool included selection bias, study design, confounders, blinding, data collection methods, and withdrawals and drop-outs. Each domain was rated as 1 (strong), 2 (moderate) or 3 (weak). A global score was calculated, in which strong=no weak ratings, moderate=1 weak rating, and weak=two or more weak ratings.

Data synthesis

A narrative synthesis was conducted as a formal metaanalysis appeared not possible at an early stage due to the heterogeneity of study designs and outcomes reported. Overall, the synthesis was split up by the micro- and meso-/macro-level. The impact of PROMs interventions was assessed by four possible determinants for increased effectiveness. The applications were categorized into mechanisms applied based on commonalities between PROMs interventions. Subsequently, we captured a broader perspective by determining the impact of PROMs interventions by the medical domain, the type of PROM used in the intervention, and by the separate outcome dimensions used to measure the effect of the intervention. For the latter, we decided to only present those which were measured in at least 3 studies. We discerned studies which used the same PROM outcome as in the intervention from studies which (only) used different outcomes. Finally, for each determinant and outcome dimension, the average quality of studies was calculated.

Results

The PRISMA diagram depicting the selection process is presented in Fig. 1. A total of 18,652 records were identified. After removing duplicates, 11,121 records were screened at title-abstract level, of which 159 were screened at full-text; 57 records were found to be eligible for inclusion [17, 19, 28, 34–88]. Through reference tracking another 21 records were identified [17, 89–108], leading to a total of 78 included studies. Two studies presented outcomes in two separate publications; these were combined resulting in 76 unique studies [17, 74, 75, 87].

Study characteristics

An overview of study characteristics, PROMs used, overall study impact and quality is presented in Table 1 (micro-level) and Table 2 (meso-/macro-level). Below we shortly describe the included studies: for a more detailed description of study characteristics refer to Supplementary Material 2, and for extended tables of study characteristics, quality assessment and outcomes extracted refer to Supplementary Material 3.

Micro-level

Sixty-eight out of 76 studies provided evidence on the use of PROMs at the micro-level [17, 19, 28, 34–36, 38–44, 46–48, 50–56, 58, 59, 62–71, 73–93, 95–108]. Most studies were conducted in the United States (n=32), and were in the medical domains primary care (n=17), oncology (n=19), gastroenterology (n=5) and

orthopedic (trauma) surgery (n=6). Fifty-five studies used a disease-specific instrument in their intervention, 3 used a generic instrument and 10 a combination. Sixteen studies were of strong quality, 31 were of moderate quality and 21 were of weak quality.

Macro-level

Eight out of 76 studies provided evidence of the use of PROMs at the macro-level [37, 45, 49, 57, 60, 61, 72, 94], and no studies were found at the meso-level. Studies were conducted in various countries. Most studies were conducted in surgical fields (n=7), of which 3 in both non-surgical and surgical fields; the eighth study was conducted in primary care. Five studies used a disease-specific PROM, 1 used a generic PROM, and 2 used a combination. Four studies were rated as moderate quality, while the other 4 were rated as weak quality.

Impact by determinants and outcome dimensions

Outcome of PROMs interventions by determinants are summarized in Table 3 (micro-level) and Table 4 (meso-/ macro-level). Table 5 shows the impact by outcome dimensions. The quality of studies for each determinant generally indicated "moderate" quality, both at the microand meso-/macro-level; the exception is highlighted. Six mechanisms were identified at the micro-level, and 3 at the meso-/macro-level.

Impact by mechanism *Micro-level*

Feedback of PROMs to patient One of 5 studies employing feedback of PROMs to patients fed back (raw) scores directly [54], 3 included a graphical display of PROMs scores [55, 78, 85], and 1 combined a narrative report with a graphical display [43]. Studies were conducted in various domains. One (20%) study conducted in headcancer patients fed back data from a comprehensive inventory of disease-related symptoms and found an improved overall outcome, driven by improved symptoms (pain and activity), mental and physical functioning [54].

Feedback of PROMs to provider Two of the 18 studies employing feedback of PROMs to providers used (raw) scores in their report [79, 90], 4 included a narrative report [52, 53, 73, 93], 8 included a graphical display [17, 36, 44, 47, 48, 84, 91, 92], and 3 combined a narrative report with a graphical display [34, 41, 89]. Overall, nine (53%) studies found an improvement in outcome [17, 34, 47, 53, 73, 84, 89, 90].



Fig. 1 PRISMA flowchart of study selection

When looking at the information collected, 14 of 18 studies fed back PROMs to patients which covered disease-specific information such as hip functioning, cancer-related, or gastrointestinal symptoms [17, 34, 36, 41, 47, 53, 73, 79, 84, 89–93]. Of these 14 studies, 9 (64%) found an improvement in outcome [17, 34, 47, 53, 73, 79, 84, 89, 90]. Most studies pertained to cancer-related symptoms (n=8) of which 5 (63%) reported

an improvement via various outcome dimensions, including reduced emergency department (ED) visits or readmissions (n=2), improved physical, mental and social functioning (n=1), symptoms (depression and cancer-related) (n=1) or experience with care (n=1) [17, 47, 79, 84, 89]. The remaining 4 studies fed back PROMs to the provider pertaining to general HRQoL and/or pain, and found no improvement in outcome [44, 48, 52, 55].

\sim
a.
F
9
$\underline{-}$
g
÷
ĕ
E
\geq
\Box
D.
ť
ō
S
Ð
\geq
-
2
<u> </u>
Ε
a)
-
t
σ
S
ŝti
÷.
te
Я
E.
Ĕ
U
R
ĭ
St
_
5
<u> </u>
-9

Author (year)	Domain	Patients	Number of patients (I = Intervention, C = Control)	PROM(s) used (GEN = generic, DS = disease-specific)	Overall impact	Quality of study
Decision-aid						
Jayakumar (2021)	Orthopedics	Considering knee replacement for OA	l:69; C:60	KOOS-JR (DS), PROMIS Global 10 (GEN)	+	Moderate
Bansback (2022)	Orthopedics	Considering knee replacement for OA	l:82; C:81	EQ-5D-5L (GEN)	+	Weak
Volkmann (2015)	Rheumatology	Considering knee replacement for OA	I:111; C:NA	WOMAC (DS)	+	Moderate
Feedback to patient						
Ngo (2022)	Gyneacology	Pregnant women	l:89; C:103	PUQE (DS)	٤	Moderate
Rogers (2021)	Oncology	Head-neck cancer, curatively treated	l:140; C:148	PCI Head Neck Cancer (DS)	+	Moderate
Steele Gray (2021)	Primary care	Elderly	l:23; C:21	PROMIS Global Health Scale (GEN), Pain Interference (DS), HAQ (GEN)	٤	Weak
Berdal (2023)	Rehabilitation	Rheumatic or musculoskeletal diseases	l:168;C:206	PSFS (DS), EQ-5D-5L (GEN), EQ VAS (GEN), 30 s sit-to-stand test	٤	Moderate
Gossec (2016) Feedback to provider	Rheumatology	Reumatoid arthritis	l:159; C:161	RAPID-3/HAQ2 (GEN), RAID (DS), symptoms as free text	٤	Moderate
Hadjistavropoulos (2009)	Community care	Elderly, with complex medical problems	l:88; C.56	Pain Assessment Battery (DS), GPM (DS), GDS short form (DS), pain drawing	٤	Weak
Almario (2016)	Gastroenterology	Gastrointestinal disease	I:217; C:154	PROMIS Gastrointestinal (DS)	ł	Moderate
Kjaer (2016)	Oncology	Head-neck cancer, finished treatment	l:132; C:134	HADS (DS), symptoms relevant to head-neck cancer survivors	+	Weak
Velikova (2004)	Oncology	Any cancer, commencing treatment	l:144, 70; C:72	EORTC-QLQ-C30 (DS), and HADS (DS)	+	Weak
Detmar (2002)	Oncology	Any cancer, undergoing palliative chemotherapy	I:114; C:100	EORTC-QLQ-C30 (DS)	ł	Weak
Rosenbloom (2007)	Oncology	Advanced breast, lung or colorectal cancer	l:71, 73; C:69	FACT-G (DS), including relevant cancer-type subscale	ł	Strong
Hilarius (2008)	Oncology	Any cancer, undergoing adjuvant or palliative chemotherapy	E111; C:108	EORTC-QLQ-C30 (DS), including relevant cancer-type (breast (QLQ-BR23), colorectal (QLQ-CR38), or lung cancer (QLQ-LC13))	ł	Strong
Barbera (2015)	Oncology	Stage I–III breast cancer, received adjuvant chemotherapy	I:2541; C:5818	ESAS (DS)	+	Strong
Patel (2022)	Oncology	Advanced cancer	1:64; C:64	ESAS (DS)	+	Strong
Skovlund (2021)	Oncology	Metastic melanoma	l:137; C:142	EORTC-QLQ-C30 (DS), HADS (DS)	+	Weak

Table 1 (continued)						
Author (year)	Domain	Patients	Number of patients (I = Intervention, C = Control)	PROM(s) used (GEN = generic, DS = disease-specific)	Overall impact	Quality of study
Ackermans (2017)	Orthopedics	Hip or knee OA	l:72; C:70	HOOS-PS/KOOS-PS (DS), NRS Pain (DS)	+	Strong
Holm (2020)	Orthopedics	Hip or knee OA	l:6245; C:N/A	NRS pain (DS), HOOS/KOOS (DS), EQ-5D (GEN), PSFS (DS), physical activity (DS), OA-QI (DS), ASES (DS)	+	Strong
De Wit (2008)	Pediatrics	Diabetes	l:46; C:45	PedsQL (GEN), Generic and Diabetes module (DS)	+	Moderate
Fihn (2004)	Primary care	Veterans	l:5801; C:3218	SF-36 (GEN), Seattle Outpatient Satisfaction Questionnaire (PREM), 1 of 6 disease-specific questionnaires (Seattle Angina, Seattle Obstructive Lung Disease, Drinking Practices, Seattle Diabetes, Seattle Hypertension, HSCL)	ł	Weak
Kroenke (2018)	Primary care	Visiting for any reason	l:151; C:149	SPADE symptoms (GEN)	٤	Moderate
Reiber (2004)	Primary care	Veterans	l:3701; C:2020	Seattle Diabetes Questionnaire, SF-36 (GEN), SOSQ (PREM)	٤	Weak
Richardson (2008)	Primary care	Elderly	l:134; C:131	Self Report Task Modification and Disability Scale (DS), HUI (GEN), SF-36 (GEN), several physical and functional performance measures	+	Moderate
Santana (2010) Monitoring	Pulmonary medicine	Pre- and post-lung transplant	l:108; C:105	HUI2/3 (GEN)	٤	Moderate
Davidson (2010)	Cardiology	Myocardial infarction	1:80; C:77	PHQ-9 (DS)	+	Moderate
de Jong (2017)	Gastroenterology	IBS	I:465; C:444	Monitor IBD At Home (DS)	+	Moderate
Berinstein (2022)	Gastroenterology	IBS	l:100; C:105	CD-PRO (DS), UC-PRO (DS)	+	Strong
Cross (2019)	Gastroenterology	IBS	l:115, 116; C:117	HBI (DS), SSCAI (DS)	+	Moderate
Pooni (2023)	General surgery	Colorectal surgery	l:128; C:125	QoR-15 (DS)	+	Strong
Girgis (2009)	Oncology	Non-localized breast and colorectal cancer	l:120,119; C:117	HADS (DS), EORTC-QLQ-C30 (DS), SCNS short form (DS), selected items from the NA-ACP (DS)	٤	Moderate
Cooley (2022)	Oncology	Lung cancer	I:89; C:91	PHQ-9 (DS), SDS (DS), FACT-G (DS), AUDIT (DS), MSAS (DS)	ł	Weak
Price (2023)	Oncology	Cancer patients with depression	I:165; C:NA	PHQ-9 (DS) and GAD-2 (DS)	+	Weak

Table 1 (continued)						
Author (year)	Domain	Patients	Number of patients (I = Intervention, C = Control)	PROM(s) used (GEN = generic, DS = disease-specific)	Overall impact	Quality of study
Livanainen (2023)	Oncology	Colorectal cancer, undergoing chemotherapy	l:36; C:35	NCTCAE (DS)	ĩ	Weak
Basch (2016)	Oncology	Metastatic breast, genitourinary, gynecologic or lung cancer	l:441; C:325	NCTCAE (DS)	+	Moderate
Sharpe (2014)	Oncology	Cancer patients with depression	l:253; C:247	PHQ-9 (DS)	+	Weak
Patel (2020)	Oncology	Hematologic or advanced stage cancer	l:186; C:102	Edmonton Symptom Assessment Scale (DS)	+	Strong
Maguire (2021)	Oncology	Non-metastatic breast, colorectal, (non-)Hodgkin's cancer, undergoing chemotherapy	l:415, C.414	DCTAQ (DS)	+	Strong
Epstein (2007)	Pediatrics	ADHD	l:162; C:215	CPRS (DS), CTRS (DS)	٤	Moderate
Dobscha (2006)	Primary care	Moderate to severe depression	l:189; C:186	PHQ-9 (DS)	+	Strong
Balestrieri (2020)	Primary care	Depression	l:66; C:32	PHQ-9 (DS) and IDS-SR (DS)	+	Moderate
Dhingra (2021)	Primary care	Chronic pain	l:256; C:272	FPS (DS), SBIRT (DS), PHQ-2 (DS), BPI short form (DS), PROMIS Pain Interference short form (DS)	ł	Weak
Katon (1996)	Primary care	Depression	l:77; C:76	BDI short form (DS)	+	Moderate
Katzelnick (2000)	Primary care	Depression	l:218; C:189	HAM-D (DS)	+	Moderate
Unützer (2002)	Primary care	Depression	l:218; C:189	PHQ-9 (DS)	+	Weak
Simon (2011)	Primary care	Depression	l:106; C:102	PHQ-9 (DS)	+	Weak
Katon (2004)	Primary care	Diabetes patients, with depression	l:164; C:165	PHQ-9 (DS)	+	Weak
Carola Pèrez (2021)	Psychiatry	Depression	I:84; C:83	PHQ-9 (DS)	ł	Moderate
Tirelli (2020)	Reumatology	Juvenile Idiopathic Arthritis, subtype Enthesitis-Related Arthritis	l:54; C:98, 51	JADAS (DS), PROMIS physical function (DS), NRS pain (DS)	٤	Moderate
Buckley (2020)	Rheumatology	Juvenile Idiopathic Arthritis	I:97; C:NA	JADAS (DS), PROMIS physical function (DS), NRS pain (DS)	+	Moderate
No feedback						
Baker (2023)	Dermatology	Eczema	l:147; C:149	POEM (DS)	+	Moderate
McCambridge (2007)	General public	University students	I:217; C:204	AUDIT (DS)	+	Moderate
Jaensson (2017)	General surgery	Various types of day surgery	l:513; C:514	SwQoR (DS)	+	Strong
Screening						
Frasure-Smith (1997)	Cardiology	Myocardial infarction	l:692; C:684	GHQ (DS)	Z	Strong
Kronish (2020)	Cardiology	Myocardial infarction	l:499, 501; C:500	PHQ-8 (DS)	٤	Weak
Allen (2011)	Community care	War veterans	1:97; C:NA	dass (ds), audit (ds)	+	Weak
van der Zee-van den Berg (2017)	Midwife care	Mothers who just gave birth	l:1843; C:1246	EPDS (DS)	+	Moderate

Table 1 (continued)						
Author (year)	Domain	Patients	Number of patients (I = Intervention, C = Control)	PROM(s) used (GEN = generic, DS = disease-specific)	Overall impact	Quality of study
Ferrell (2021)	Oncology	Palliative cancer, undergoing a therapeutic clinical trial	l:239; C:240	Distress Thermometer (DS), FACT-G (DS)	+	Weak
Shyu (2013)	Orthopedic and trauma surgery	Hip fracture patients	l:99; C:101, 99	GDS short form (DS)	+	Strong
Mallen (2017)	Primary care	Hip or knee osteoarthritis patients	l:1339; C:703	GAD-2 (DS), PHQ-2 (DS), NRS pain (DS)	ı	Moderate
Fortmann (2020)	Primary care	Diabetes patients	l:236;C:239	PHQ-9 (DS)	+	Moderate
Rollman (2021)	Cardiology	Heart failure	l:251; C:252, 126	PHQ-2 and -9 (DS)	+	Moderate
Regueiro (2019)	Gastroenterology	IBS	I:322; C:NA	Harvey-Bradshaw Index (DS) for CD or Ulcerative Colitis Activity Index for UC, Short Inflammatory Bowel Disease questionnaire (DS), PHQ-9 (DS), GAD-7 (DS)	+	Moderate
Howell (2020)	Oncology	Various types of cancer	l:13,260, 10,875; C.57,594, 48,068	ESAS (DS), DART (DS), BPI (DS), CFS (DS), GAD (DS), PHQ (DS), occasionally the CPC (DS) or SDI (DS)	+	Strong
Wylde (2022)	Orthopedics	Received knee replacement surgery	l:242; C:121	BPI (DS), HADS (DS), PainDETECT (DS), Douleur Neuropathique 4 (DS)	+	Weak
Buxton (2012)	Primary care	Low income and uninsured	I:36; C:81	SF-12 (GEN), AUDIT (DS), PHQ-9 (DS)	ł	Moderate
Wu (2018)	Primary care	Diabetes type 2	I:366, 380; C:341	PHQ -2 and/or -9 (DS)	+	Strong
Studies are categorized according ADHD Attention Deficit Hyperactiv Cancer Fatigue Scale, COOP Primar Tool, DASS Depression Anxiety Stre questionnaire, EPDS Edinburgh Po: Pain Scale, GAD Generalized Anxiet Rating Scale, HAO Health Assessme Rating Scale, HAO Health Assessme Replacement, KOOS-PS Knee disabi Nark, IBS Inflammatory Bowel Dise Replacement, KOOS-PS Knee disabi National Cancer Institute's Commo Concerns Inventory, PedsOL Pediat measures, PROMIS Patient-Reporte RAID Reumatoid Arhritis Impact of Social Difficulties Inventory, 205 Sy	to the mechanism applied. Additional st ty Disorder, ASES Arthritis Self-Efficacy 5 (Care Cooperative Information Project, ss Scales, DCTAQ Daily Chemotherapy Ti thatal Depression Scale, EQ-5D-5L Five- y Disorder, GDS Geriatric Depression Sc in tOuestionnaire, HBI Harvey-Bradshaw ase, IDS-5R Inventory of Depressive Sym lity and Osteoarthritis Outcome Score- or Terminology Criteria for Adverse Even Terminology Criteria for Adverse Even Terminology Criteria for Adverse Even of Outcomes Measurement Information Disease, RAPID-3/HAQ2 RAPID3 Health. Ampton Distease Scale, 57-36 Short Form	tudy characteristics, full quality assess Scale, <i>AUDIT</i> Alcohol Use Disorders Idd <i>CPC</i> Canadian Problem Checklist, <i>CPI</i> oxicity Self-Assessment Questionnalin level version of the EQ-5D, ESAS Edmu- ale, <i>GHQ</i> General Health Questionnal undex, <i>HOOS-PS</i> Hip disability and O. ptomatology – Self Rated, <i>JADAS</i> Clir ptomatology – Self Rated, <i>JADAS</i> Clir thealth Questionnaire, <i>POEM</i> Patient System, <i>PSFS</i> Patient Specific Functio Assessment Questionnaire, <i>SBIRT</i> Scre to the health Survey, <i>SOS</i> 0 Seattle Outpati	sment and all extracted outcomes can lentification Test, BD/ Beck Depression Ir BS Conners Parent Baring Scale, CTRS Constrainers Parent Baring Scale, CTRS Conners Parent Baring Scale, CTRS Context Sconence, FADS Here, GPM Geriatric Pain Measure, HADS Here, GPM Geriatric Pain Measure, HADS Here, GPM Geriatric Pain Measure, HADS Here, GPM Geriatric Pain Measure, PACP Nee, teo and Scale, Nu-ACP Nee, Toton Seesment Score – Physical ical Juvenile Arthritis Disease Activity 5 and ical Juvenile Arthritis Disease Activity 2 for the Activity 5 and Contented Eczema Measure, PREM spatianal Scale, PUGE Pregnancy Unique Quation Brief Intervention, and Referral tert Satisfaction Question	be found in Supplementary Material 3, aventory, BPI Brief Pain Inventory, CD-P onners Teachers Rating Scale, DART Dis itzation for Research and Treatment of ACT-G Functional Assessment of Cance lospital Anxiety and Depression Scale, Eunction, HSCL Hopkins Symptom Che Scone, KOOS-JR Kneel Injury and Osteoa eds Assessment for Advanced Cancer P foundite, OH5 Oxford Hip Scone, OKS O ionnaite, OH5 Oxford Hip Scone, OKS O ent-reported experience measures, PR infification of Emesis, OoR-15 Quality o to Treatment questionnaire, SCNS Supp Supp.	Table 1 to PRO Crohn's PRO Crohn's stress Assess Cancer ques r Therapy r Therapy HUM Phan AMA-D Han cklist, HUI r thritis Out r thritis Out vertice Care wenergyff, wenergyff,	B Disease PRO, CF5 sment Response lity of life -General, FP5 Faces milton Depression Health Utilities Index come Score for Joint stionnaire, <i>NCTCAE</i> stronaire, <i>NCTCAE</i> stronaire, <i>NCTCAE</i> stronaire, <i>NCTCAE</i> theore a stronaire, <i>NCTCAE</i> stronaire, <i>NCTCAE</i> stronai
כוונוורפו בטוונוז ארנועויל ווומבע, שייעני	ת אששטואון בעמוונע טו הבנטעבוע שניבי שנ	שכוא גראט עונפו בחטי איז איז איז איז איז איז איז איז איז אי	מן אוומוטטער טרמוב, איטואואר ערטובווו טוו	טבוכט כבוווניוטוווע ובונעועוועווע פוועט	יווו כחו וו ח ומ	xər

Table 2 Study characteristics at the meso-/macro-level (sorted by medical domain)

Author (year)	Domain	Patients	Number of patients	PROM(s) used (GEN = generic, DS = disease-specific)	Overall impact	Quality of study
Benchmarking						
Boyce (2015)	Orthopedics	Receiving primary hip arthroplasty	l:230; C:228	OHS (DS)	~	Weak
Varagunam (2014)	Orthopedics and general surgery	Receiving hip or knee arthroplasty, varicose vein or inguinal hernia surgery	l: 7 k-30 k; C:NA	EQ-5D-5L (GEN), EQ VAS (GEN), 1 of 3 disease-specific questionnaires (OHS, OKS, AVVQ)	~	Moderate
Weingarten (2000)	Primary care	Elderly	l:541; C:543	Dartmouth COOP (GEN)	2	Moderate
Kumar (2021)	Urology	Undergoing prostate surgery for cancer	l:212; C:210	Selected items (continence, sexual function) from the EPIC (DS)	-	Weak
In-depth analysis of d	lata					
Haller (2011)	Internal and surgical departments	Wards	l:1237; C:1113	IPO questionnaire (DS), NRS pain	+	Moderate
Zaslansky (2019)	Surgical departments	Wards	I:? C:?	IPO questionnaire (DS), NRS pain	+	Weak
Garduño-López (2021)	Surgical departments	Wards	I:? C:?	IPO questionnaire (DS), NRS pain	+	Weak
PDCA-cycle						
Partridge (2016)	Orthopedics	Patient receiving total knee arthroplasty	l:827, C:441	OKS (DS), EQ-5D-3L (GEN)	+	Moderate

Studies are categorized according to the mechanism applied. Additional study characteristics, full quality assessment and all extracted outcomes can be found in Supplementary Material 3, Table 1 to 3

AVVQ Aberdeen Varicose Vein Questionnaire, COOP Primary Care Cooperative Information Project, EPIC Expanded Prostate Cancer Index Composite, EQ-5D-3L Threelevel version of the EQ-5D, EQ-5D-5L Five-level version of the EQ-5D, IPO International Pain Outcomes Questionnaire, NRS Numerical Rating Scale, OHS Oxford Hip Score, OKS Oxford Knee Score, VAS Visual Analogue Scale

Using PROMs to screen for disease or symptoms Seven studies out of 14 used PROMs to screen for depression [28, 35, 50, 56, 71, 98, 102], and 1 study for oncological symptoms [70], to initiate treatment or a care pathway. Of these, five (63%) studies observed an improved outcome driven by improved symptoms (depression, stress or anxiety) (n=4), improved mental (n=2), social (n=2), and physical functioning (n=1), and reduced ED visits and readmissions (n=1) [28, 35, 56, 70, 71]. One study found an outcome deterioration via worsened pain symptoms [50].

Six studies combined the screening for depression with follow-up monitoring to evaluate whether the treatment works, and potentially adjust if treatment was ineffective [38, 59, 74, 83, 88, 105]. Of these, three also incorporated disease-specific information: knee functioning [88], cancer-related [74], and gastro-intestinal symptoms [105]. Five (83%) out of 6 studies found improved outcome particularly via improved symptoms (depression and anxiety) (n=4) and reduced ED visits (n=2) [59, 74,

83, 88, 105]. Two of three disease-specific symptoms also improved, except for oncological symptoms [74].

Using PROMs to monitor symptoms Twelve out of 25 studies used PROMs to identify patients under treatment exceeding predefined thresholds of symptoms and linked these to treatment changes, increased monitoring or care pathways [39, 63, 66, 67, 81, 86, 95, 97, 100, 103, 107, 108]; 10 (83%) found an improved outcome [39, 63, 66, 81, 95, 97, 100, 103, 107, 108]. Seven studies also used PROMs monitor treatment but did not explicitly mention the use of predefined algorithms [40, 42, 69, 82, 99, 101, 104]; 4 (57%) reported an improvement [82, 99, 101, 104]. Six studies incorporated PROMs into the clinical pathway and sent out alerts upon exceeding a threshold without specific guidance to the provider [64, 68, 76, 80, 96, 106], 1 of these also used PROMs to monitor treatment response [106]; three (50%) found an improved outcome [64, 96, 106].

When looking at the information collected, 13 out of 25 studies used PROMs to monitor existing depression symptoms [42, 63, 68, 69, 80, 82, 97, 99–101, 106–108].

		Number of studies	Improvement (%)	Quality of studies (average)
Mechanism	Feedback to patient	5	1 (20)	1.8
	Feedback to provider	18	9 (50)	1.9
	Screening	14	10 (71)*	2.0
	Monitoring	25	18 (72)	1.9
	No feedback	3	3 (100)	2.3
	Decision-aid	3	3 (100)	1.7
Medical domain	Cardiology	4	2 (50)	2.0
	Community care	2	1 (50)	1.0
	Dermatology	1	1 (100)	2.0
	Gastroenterology	5	4 (80)	2.2
	General public	1	1 (100)	2.0
	General surgery	2	2 (100)	3.0
	Gyneacology	1	0 (0)	2.0
	Midwife care	1	1 (100)	2.0
	Oncology	19	13 (68)	1.9
	Orthopedics/trauma surgery	6	6 (100)	2.2
	Pediatrics	2	1 (50)	2.0
	Primary care	17	10 (59)*	1.7
	Psychiatry	1	0 (0)	2.0
	Pulmonary medicine	1	0 (0)	2.0
	Rehabilitation	1	0 (0)	2.0
	Rheumatology	4	2 (50)	2.0
Type of PROM	Disease-specific	55	39 (71)*	2.0
	Generic	4	1 (25)	1.5
	Combination	9	4 (44)	1.9

Table 3 Overall impact by determinants at the micro-level

* One study showed a deterioration

PROM Patient-Reported Outcome Measure

Table 4 Overall impact by determinants at the meso-/macro-level

		Number of studies	Improvement (%)	Quality of studies (average)
Mechanism	Benchmarking	4	0 (0)*	1.3
	In-depth analysis of data	3	3 (100)	1.7
	PDCA-cycle	1	1 (100)	2.0
Medical domain	Orthopedics	2	1 (50)	1.7
	Primary care	1	0 (0)	1.0
	Urology	1	0 (0)*	2.0
	Various internal and surgical departments	4	3 (75)	1.5
Type of PROM	Disease-specific	5	3 (60)*	1.4
	Generic	1	0 (0)	1.0
	Combination	2	1 (50)	2.0

* One study showed a deterioration

PROM Patient-Reported Outcome Measure

Table 5 Impact by outcome dimensions

		Micro-level			Meso-/macro-level	
	Number of studies	Improvement (%)	Quality of studies (average)	Number of studies	Improvement (%)	Quality of studies (average)
PROMs						
Functioning						
Physical	29	12 (41)	1.9	2	0	1.0
Mental	25	8 (32)	1.8	1	0	1.0
Social	16	6 (38)	1.8	1	0	1.0
HRQoL	29	11 (38)	1.8	5	2 (40)	1.6
Role limitations						
Physical	5	1 (17)	2.0	-	-	-
Emotional	5	0	2.0	-	-	-
General health perceptions	8	6 (75)	2.4	1	0	1.0
Symptoms combined	46	26 (57)**	1.9	5	3 (60)**	1.6
Depression	25	14 (56)	1.8	2	0	2.5
Anxiety	14	6 (43)	2.0	2	1 (50)	2.5
Alcohol use/disorder	3	1 (33)	1.7	-	-	-
Pain	17	7 (42)**	1.9	5	2 (40)	1.6
Vitality/fatigue	5	3 (60)	1.8	1	0	3.0
Nausea	4	0	2.3	2	1 (50)	1.0
Decision-conflict and readiness	4	3 (75)		-	-	-
PREMs						
Satisfaction	23	10 (43)	2.0	-	-	-
Patient-physician relationship	5	1 (20)	2.0	-	-	-
Experience with care	7	4 (57)	2.0	-	-	-
Supportive needs	3	1 (33)	2.7	-	-	-
Patient-activation	7	4 (57)	2.1	-	-	-
Physician awareness of HRQoL	2	0	2.0	-	-	-
Clinical outcomes						
Complications*	8	1 (13)	1.8	2	0	1.5
(Re)admissions	17	5 (29)	2.2	-	-	-
Emergency department visits	12	7 (58)	2.4	-	-	-
Survival	5	0	2.0	-	-	-
Lab values	4	2 (50)	2.0	-	-	-
Outcome same as PROM used in intervention	32	18 (56)**	1.8	7	3 (42)**	1.4
Outcome not the same as PROM used in intervention	36	26 (72)	2.0	1	1 (100)	2.0

*Complications also vary by domain and intervention, e.g., a bleed in myocardial infarction patients

**One study showed a deterioration

PROM Patient-Reported Outcome Measure; PREM Patient-Reported Experience Measure, HRQoL Health-Related Quality of Life

Of these, 10 (77%) found an improved outcome, mostly driven by improved depression symptoms (n=9) and satisfaction (n=5) [63, 69, 82, 97, 99–101, 106–108]. Five studies used PROMs to monitor cancer-related symptoms [64, 67, 76, 103, 104], of which 3 (60%) found various improved outcomes including HRQoL, physical and mental functioning, and satisfaction [64, 103, 104]. Three studies monitored gastro-intestinal symptoms in patients

with inflammatory bowel disease and all (100%) found reduced readmissions (n=2) and improved HRQoL (n=1) [39, 66, 96]. The remaining 4 studies were conducted in various domains [40, 81, 86, 95], of which two showed improved outcomes. The first monitored surgical recovery in colorectal surgery patients and found improved perception of general health, anxiety and satisfaction. The other used PROMs to guide treatment in children with juvenile idiopathic arthritis and found reduced pain and arthritis activity [81, 95].

No feedback: filling out effect of PROMs One of 3 studies tested the hypothesis of whether merely filling out alcohol abuse PROMs would reduce alcohol use by a direct measurement effect [51]. Similarly, another study collected PROMs weekly in patients with eczema without any additional interventions [62]. The third study collected PROMs daily after surgery via an app; patients could always contact their provider via the e-portal [46]. All (100%) studies reported improved outcome due to improved symptoms (depression and alcohol dependency) (n=2) and improved HRQoL (n=1).

PROMs in decision-aids In three studies a one-time PROM was used in a decision-aid along an education component to help with treatment choice (surgical vs. conservative) in patients with knee osteoarthritis [19, 58, 77]. All studies (100%) found an improvement in shared-decision making, while 1 of these only found this effect in females [58].

Meso-/macro-level

PROMs in benchmarking Three benchmarking studies used case-mix adjusted PROM scores [37, 49, 57], while the fourth used unadjusted scores [94]. Three studies presented performance reports to the provider, which included PROM scores and how they compared to peer providers [37, 49, 94]; in 2 studies complication rates were also presented [37, 49]. The other study evaluated a nationwide PROMs collection program, which provided both patients and providers the option to check providers' PROMs outcomes [57]. All studies were of weak quality, and did not find an improvement in outcome; 1 study even reported a potential worsening [49].

PROMs in in-depth analysis of data Three studies used PROM data in combination with guidelines, teaching and protocols to improve pain management in various surgical and non-surgical departments [45, 60, 72]. One of these studies also used a feedback loop by a department representative to evaluate and provide advice on the implemented initiatives [45]. The two other studies pertained to the same quality initiative aimed to reduce the pain of patients admitted to hospitals but were conducted in different developing countries/departments [60, 72]. All 3 (100%) studies found an improvement in outcome due to reduced pain (n=3) and nausea (n=2) symptoms in particular. *PROMs in PDCA-cycles* One study conducted a PDCA-cycle where they introduced an improved total knee implant and changed their surgical technique, guided by and evaluated with PROMs scores [61]: an overall improvement in outcome (HRQoL) was observed.

Impact by medical domain *Micro-level*

At the micro-level, the medical domains in which PROM interventions were conducted which seemed to be consistently associated with improved outcome were orthopedic (trauma) surgery (n=6 studies, 100% effective), gastroenterology (n=5, 80%), oncology (n=19, 68%), and primary care (n=17, 59%). Less effective seemed cardiology (n=4, 50%) and rheumatology (n=4, 50%). Limited evidence was available for other domains.

Meso-/macro-level

Interventions conducted in orthopedics, primary care, and urology were not found to be related to improved outcome. Four studies covered various internal and surgical departments, of which 3 (75%) showed improved outcome.

Impact by type of PROM used in intervention *Micro-level*

Most studies used a disease-specific PROM, which showed the highest percentage of improved outcomes (n=55 studies, 71% effective). Generic PROMs or a combination of both showed an overall lower percentage (n=13, 38%). While disease-specific PROMs were used in all mechanisms, generic PROMs were used in studies employing the "feedback" mechanism (n=10), "decision-aids" (n=2), and once (combined with a disease-specific PROM) in "screening".

Meso-/macro-level

According to the type of PROM (disease-specific vs. generic) no specific pattern was observed.

Impact by outcome dimensions

Micro-level

In this section, we describe the impact of the PROMsinterventions on the outcome dimensions (PROMs, PREMs or clinical outcomes), regardless of the mechanism or other determinants.

Regarding PROMs, studies often showed an improvement in general health perceptions (n=8 studies, 75% effective), decision-readiness and conflict (n=4, 75%) and symptoms overall (n=46, 57%). Particularly depression was evaluated often (n=25), and improved

in 57% of studies. The percentage decreased for HRQoL (n=29, 38%) and physical and mental functioning domains.

Regarding PREMs, satisfaction was most often studied (n=23), and improved in less than half of studies (43%). Patient-activation and experience with care tended to improve slightly more often (n=7, 57%, for both outcomes).

As for clinical outcomes, twelve studies analyzed emergency department visits, of which 58% found an improvement. Fewer studies observed a positive effect on complications (n=8, 13%) and (re)admissions (n=17, 29%), and no studies observed an effect on survival (n=5, 0%).

Studies which used a different outcome than the PROM in the intervention more often had an improved overall outcome (n = 36, 72%), compared to those which did not (n = 32, 56%).

Meso-/macro-level

With regard to PROMs, symptoms showed improved most often, which mostly pertained to pain (n=5, 60%). HRQoL was also measured in 5 studies, however, improved in less studies (40%). Other domains and outcomes were studied in only a few studies, and showed no improvement.

Discussion

In this systematic review, evidence on the use of PROMs to improve patient-related outcomes at the micro- (68 studies) and meso-/macro- [8] levels was collected and analyzed. Moreover, determinants for increased effectiveness were elucidated.

At the micro-level, 44% of studies employing direct feedback of PROMs to the provider and/or patient resulted in improved patient outcomes, which is in line with previous reviews [8–15]. A contemporary development was to use PROMs to screen for disease or to monitor existing disease. These studies linked the PROMs scores to care pathways or treatment adaptations, and approximately 70% of studies found improved outcomes. This approach was particularly effective for depression, oncological and gastroenterological disease. A novel application was to use PROMs to inform patients considering knee arthroplasty, which generally resulted in improved decision-quality. At the meso-/macro-level, current evidence does not support using PROMs in benchmarking. The scarce evidence available suggests, however, that PROMs might be of value in an in-depth analysis of the performance of departments and hospitals and PDCA-cycles. At both the micro- and meso-/ macro-level, studies more often employed disease-specific PROMs, which – in comparison with studies which employed generic PROMs – found improved outcomes more often.

The evidence at all levels was of moderate quality at best, which raises concerns regarding the validity of the findings.

Micro-level

Providing feedback on the PROM scores to patients or providers is generally thought to benefit outcomes via improved patient-healthcare professional communication and identification of problematic symptoms [16]. This application is often used in patients with chronic disease who have multiple visits to their doctor, which in our review included diabetes, gastrointestinal disease, oncology, orthopedics, transplantation care; most evidence was available for oncology [8, 27]. For example, two studies applied a tailored symptom inventory for head-neck cancer patients and found a positive impact on PROMs [47, 54]. The effectiveness may be because this group presumably experiences a number of severe physical symptoms (e.g., problems with swallowing) which, if timely detected, are sensitive to treatment.

The application of PROMs to improve patient outcomes seems particularly effective if a deviation from the acceptable threshold occurs and can be linked to a recognizable action by the clinician, such as referral or treatment adaptation. This mechanism was effective in several studies in the medical domains, including depression, oncology and gastrointestinal care. For example, monitoring patients with diagnosed diseases such as inflammatory bowel disease or screening for disease with an expected high burden in the studied population such as post-partum depression may be beneficial [28, 39]. The purpose and goal of the tool may be clearer for both patient and provider, which could increase its effectiveness.

Various reasons may underlie decreased effectiveness of PROM-interventions. Firstly, a general trend was observed that studies utilizing generic PROMs found less positive effect overall, and these studies mostly did not link a generic PROM to a care pathway (such as "screening" or "monitoring"). Generic PROMs may provide insufficient insight into treatable or modifiable factors related to the studied population. However, it should be noted, one of the identified decisionaids successfully employed only a generic measure in patients considering knee arthroplasty [19]. Combined, we believe this underlines the fact that the choice of PROM in the intervention should be driven by the intended use. Secondly, the measured outcome may play a role: PROM interventions tended to have a more pronounced impact on general health perceptions and symptom burden, but less so on certain outcomes such as HRQoL in general or survival. Other reasons for failure may include patients' resistance to discussing symptoms, time constraints in clinical practice and lack of provider continuity, and implementation hurdles through lack of knowledge [16].

The evaluation of interventions based on systematic PROM feedback appears to be a challenge. Firstly, the definition of 'control' treatment: about a third of the studies collected PROMs in the control group, unconnected to feedback or another intervention. This may decrease the difference as the collection of PROMs itself may induce beneficial effects as observed in 3 studies [46, 51, 62]. These findings suggest a Hawthornelike effect through the completion of PROMs alone [51, 109]. The patient's self-knowledge and awareness are increased, and filling out the questionnaire may increase their empowerment to take a more active role in their healthcare [34]. We expected this effect to be relatively limited, as approximately half of studies used a different outcome measure than the PROM in the intervention and generally found an improvement. Secondly, most studies did not measure intervention compliance making it impossible to know to what extent (and how) patients or providers used the PROM interventions. Thirdly, PROMs are generally part of a more complex intervention with multiple facets (e.g., patient education), and it is impossible to isolate the exact role of the PROM in the intervention. However, we believe this is also one of the key roles of PROMs in contemporary medicine; they can enhance interventions by offering important insight into patient outcomes.

Meso-/macro-level

The 4 studies which evaluated PROM benchmarking did not find a benefit. Multiple reasons for the intervention not being successful have been suggested. Boyce et al. noted that PROMs have not been developed nor validated as performance measures, and the choice of PROM may play a role in the usability of the provided feedback [37]. It is possible that inter-provider comparisons do not inherently motivate professionals to initiate additional audits and research activities or professionals may lack the knowledge to undertake such initiatives. The included studies do not describe how the data was (or wasn't) used in a feedback process of change. Kumar et al. suggested that further improvement might be prevented when the quality of care is already high [49]. The quality of the benchmarking process is also dependent on adequate case-mix variable selection, which is time-consuming and costly [110, 111]. A lack of educational support could also play a role, and it may be useful to provide examples of successes and failures with using PROMs data [112]. Finally, aggregated PROMs are used extensively in research aimed at improving quality care through, e.g. identifying subgroups at risk for poorer outcomes. These studies presumably have a large impact on national clinical guidelines, however, to our knowledge, the impact is hardly reported in peer-reviewed literature. The same applies to quality benchmarking under the supervision of professional organisations: this information is discussed with hospital groups and individuals but is generally not published.

Some examples, however, were found for in-depth analysis and PDCA-cycles with the aim to initiate quality improvements. A PDCA-cycle provides a structured and iterative approach to test changes aimed at improving the quality of systems [113]. Four studies were found that exploited these types of methods using PROMs data, all finding a benefit on patient outcomes. Zaslansky et al. suggested that the success could be attributable to the relatively low starting performance of partaking departments [60]. A commonality among these studies is the clear definition of the goal, an action plan, and feedback on the intervention along the way; all potential items which might facilitate the success of a quality improvement initiative, also highlighted by a Cochrane review [114].

Strengths and limitations

The major strength of this review is the broad search strategy, including the added value of PROMs at the micro-, meso- and macro-level. Several limitations must be acknowledged. Non-peer-reviewed literature (e.g., registry reports), which may be an important source of information on the use of PROMs as quality improvement tool, was excluded. However, this was not deemed feasible because these documents are often published in non-English languages and generally do not report clear evidence of an impact, such as a beforeafter comparison. Meta-analysis and estimating the effect sizes were not possible due to the heterogeneity of outcomes. PROM scores were variably reported as total score and/or by dimension, limiting the synthesis on the impact of PROMs-interventions by outcome dimensions.

Conclusion

This systematic review provides a comprehensive overview of novel applications of PROMs which aim improve patient outcomes, and determinants for increased effectiveness. The effectiveness appears to relate to the underlying mechanism, type of PROM used and outcome studied. At the micro-level, for example, PROMs feedback to patient or provider was positively associated with patient outcomes in approximately half of studies. Contemporary studies went a step further and linked PROMs scores to care pathways in for example depression, oncological and gastrointestinal care, which resulted in improved outcomes in a higher percentage of studies. At the meso-/macro-level evidence was limited, and evidence did not suggest a benefit of using PROMs for benchmarking. Promising applications included in-depth analysis and PDCA-cycles using PROMs data. With the increasing use of PROMs in routine clinical care, these findings may help in designing applications which truly impact patient outcomes. As the quality of studies was moderate at best raising concerns regarding the validity of findings, rigorously designed studies should be conducted on testing these applications.

Abbreviations

PROMs	Patient-rep	ported outco	ome me	asure	2S			
PDCA cycle	Plan-do-st	udy-act cycl	e					
PRISMA	Preferred Reporting Items for Systematic Reviews an							
	Meta-Analyses							
PREMs	Patient rep	orted expe	riences r	neas	ures			
HRQoL	Health-Rel	ated Quality	of Life					
ED	Emergenc	y departme	nt visits					

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12955-024-02312-4.

Supplementary Material 1 Supplementary Material 2. Supplementary Material 3.

Acknowledgements

We would like to thank dr. Wichor Bramer (biomedical information specialist at Erasmus Medical Center) for his help in developing the search strategy.

Authors' contributions

Concept and design: JB, GB, HP; Literature search: JB, AI; Quality assessment: JB, AH; Interpretation of results: JB, AI, AH; Drafting of the manuscript: JB; Critical revision of the paper for important intellectual content: all authors; Obtaining funding: JB, HP, GB; Supervision: HP, GB.

Funding

This work was funded by a PhD grant (PHD-287) provided by the EuroQol Research Foundation. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

Given the study design (systematic review) ethics approval was not required nor sought. This study was registered prospectively in PROSPERO under record 2022 CRD42022333400.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Orthopaedics and Sports Medicine, Erasmus Medical Center, University Medical Center Rotterdam, Doctor Molewaterplein 40, 3015 GD Rotterdam, The Netherlands. ²School of Public Health, University of Alberta, Edmonton, AB, Canada. ³Department of Quality and Patientcare, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, The Netherlands. ⁴EuroQol Research Foundation, Rotterdam, The Netherlands. ⁵OPEN Health Evidence & Access, Rotterdam, The Netherlands.

Received: 17 March 2024 Accepted: 30 October 2024 Published online: 26 November 2024

References

- Brooks R. EuroQol: the current state of play. Health Policy. 1996;37(1):53–72.
- Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom. 2015;84(3):167–76.
- Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Appl Health Econ Health Policy. 2017;15(2):127–37.
- Meadows KA. Patient-reported outcome measures: an overview. Br J Community Nurs. 2011;16(3):146–51.
- Murray CJ, Frenk J. A framework for assessing the performance of health systems. Bull World Health Organ. 2000;78(6):717–31.
- 6. The Organisation for Economic Co-operation and Development (OECD) Health at a Glance 2019.
- Alberta PROMs and EQ-5D Research and Support Unit (APERSU). Enhancing the Use of Patient-reported Outcome Measures (PROMs) in the Healthcare System in Alberta. Link: https://apersu.ca/wp-content/ uploads/2020/09/APERSU-PROMs-White-Paper.pdf. 2020.
- Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2013;22(9):2265–78.
- Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. Med Care. 2000;38(2):175–86.
- Gilbody SM, House AO, Sheldon T. Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome–a systematic review. Psychol Med. 2002;32(8):1345–56.
- Gilbody SM, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. BMJ. 2001;322(7283):406–9.
- Greenhalgh J, Meadows K. The effectiveness of the use of patientbased measures of health in routine practice in improving the process and outcomes of patient care: a literature review. J Eval Clin Pract. 1999;5(4):401–16.
- Ishaque S, Karnon J, Chen G, Nair R, Salter AB. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2019;28(3):567–92.
- Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2008;17(2):179–93.
- Gibbons C, Porter I, Gonçalves-Bradley DC, Stoilov S, Ricci-Cabello I, Tsangaris E, et al. Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice. The Cochrane database of systematic reviews. 2021;10(10):CD011589.
- 16. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, Black N. How do patient reported outcome measures (PROMs) support

clinician-patient communication and patient care? A realist synthesis. Journal of patient-reported outcomes. 2018;2:42.

- Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, Selby PJ. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. J Clin Oncol. 2004;22(4):714–24.
- Watson L, Link C, Qi S, Delure A, Chmielewski L, Hildebrand A, et al. Testing a modified electronic version of the Edmonton symptom assessment system-revised for remote online completion with ambulatory cancer patients in Alberta. Canada Digit Health. 2023;9:20552076231191000.
- Bansback N, Trenaman L, MacDonald KV, Durand D, Hawker G, Johnson JA, et al. An online individualised patient decision aid improves the quality of decisions in patients considering total knee arthroplasty in routine care: A randomized controlled trial. Osteoarthr Cartil Open. 2022;4(3): 100286.
- Dorr MC, van Hof KS, Jelsma JGM, Dronkers EAC, Baatenburg de Jong RJ, Offerman MPJ, de Bruijne MC. Quality improvements of healthcare trajectories by learning from aggregated patient-reported outcomes: a mixed-methods systematic literature review. Health Res Policy Syst. 2022;20(1):90.
- Greenhalgh J, Dalkin S, Gibbons E, Wright J, Valderas JM, Meads D, Black N. How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. J Health Serv Res Policy. 2018;23(1):57–65.
- Al Sayah F, Lahtinen M, Bonsel GJ, Ohinmaa A, Johnson JA. A multi-level approach for the use of routinely collected patient-reported outcome measures (PROMs) data in healthcare systems. Journal of patientreported outcomes. 2021;5(Suppl 2):98.
- 23. Ellis J. All inclusive benchmarking. J Nurs Manag. 2006;14(5):377-83.
- Willmington C, Belardi P, Murante AM, Vainieri M. The contribution of benchmarking to quality improvement in healthcare. A systematic literature review. BMC health services research. 2022;22(1):139.
- 25. Valderas JM, Fitzpatrick R, Roland M. Using health status to measure NHS performance: another step into the dark for the health reform in England. BMJ Qual Saf. 2012;21(4):352–3.
- Totten AM, Wagner J, Tiwari A, O'Haire C, Griffin J, Walker M. Closing the quality gap: revisiting the state of the science (vol. 5: public reporting as a quality improvement strategy). Evid Rep Technol Assess (Full Rep). 2012(208.5):1–645.
- Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. BMC Health Serv Res. 2013;13:211.
- van der Zee-van den Berg Al, Boere-Boonekamp MM, Groothuis-Oudshoorn CGM, MJ IJ, Haasnoot-Smallegange RME, Reijneveld SA. Post-Up Study: Postpartum Depression Screening in Well-Child Care and Maternal Outcomes. Pediatrics. 2017;140(4).
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372: n71.
- Roberts AYK, editor. Evidence-based practice manual: research and outcome measures in health and human services: New York. NY: Oxford University Press; 2004.
- Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q. 2005;83(4):691–729.
- 32. Agency for Healthcare Research and Quality R, MD. Six Domains of Healthcare Quality. Content last reviewed December 2022. [Available from: https://www.ahrq.gov/talkingquality/measures/six-domains. html.
- 33. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J Eval Clin Pract. 2012;18(1):12–8.
- Ackermans L, Hageman MG, Bos AH, Haverkamp D, Scholtes VAB, Poolman RW. Feedback to Patients About Patient-reported Outcomes Does Not Improve Empowerment or Satisfaction. Clin Orthop Relat Res. 2018;476(4):716–22.
- 35. Allen J, Annells M, Nunn R, Petrie E, Clark E, Lang L, Robins A. Evaluation of effectiveness and satisfaction outcomes of a mental health

screening and referral clinical pathway for community nursing care. J Psychiatr Ment Health Nurs. 2011;18(5):375–85.

- Almario CV, Chey WD, Khanna D, Mosadeghi S, Ahmed S, Afghani E, et al. Impact of National Institutes of Health Gastrointestinal PROMIS Measures in Clinical Practice: Results of a Multicenter Controlled Trial. Am J Gastroenterol. 2016;111(11):1546–56.
- Boyce MB, Browne JP. The effectiveness of providing peer benchmarked feedback to hip replacement surgeons based on patient-reported outcome measures-results from the PROFILE (Patient-Reported Outcomes: Feedback Interpretation and Learning Experiment) trial: a cluster randomised controlled study. BMJ Open. 2015;5(7): e008325.
- Buxton JA, Chandler-Altendorf A, Puente AE. A novel collaborative practice model for treatment of mental illness in indigent and uninsured patients. Am J Health Syst Pharm. 2012;69(12):1054–62.
- de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (mylBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet. 2017;390(10098):959–68.
- Epstein JN, Rabiner D, Johnson DE, Fitzgerald DP, Chrisman A, Erkanli A, et al. Improving attention-deficit/hyperactivity disorder treatment outcomes through use of a collaborative consultation treatment service by community-based pediatricians: a cluster randomized trial. Arch Pediatr Adolesc Med. 2007;161(9):835–40.
- Fihn SD, McDonell MB, Diehr P, Anderson SM, Bradley KA, Au DH, et al. Effects of sustained audit/feedback on self-reported health status of primary care patients. Am J Med. 2004;116(4):241–8.
- 42. Girgis A, Breen S, Stacey F, Lecathelinais C. Impact of two supportive care interventions on anxiety, depression, quality of life, and unmet needs in patients with nonlocalized breast and colorectal cancers. J Clin Oncol. 2009;27(36):6180–90.
- Gossec L, Cantagrel A, Soubrier M, Berthelot JM, Joubert JM, Combe B, et al. An e-health interactive self-assessment website (Sanoia^(®)) in rheumatoid arthritis. A 12-month randomized controlled trial in 320 patients. Joint Bone Spine. 2018;85(6):709–14.
- Hadjistavropoulos T, MacNab Y, Lints-Martindale A, Martin R, Hadjistavropoulos H. Does routine pain assessment result in better care? Pain Res Manag. 2009;14(3):211–6.
- Haller G, Agoritsas T, Luthy C, Piguet V, Griesser AC, Perneger T. Collaborative quality improvement to manage pain in acute care hospitals. Pain Med. 2011;12(1):138–47.
- Jaensson M, Dahlberg K, Eriksson M, Nilsson U. Evaluation of postoperative recovery in day surgery patients using a mobile phone application: a multicentre randomized trial. Br J Anaesth. 2017;119(5):1030–8.
- Kjaer T, Dalton SO, Andersen E, Karlsen R, Nielsen AL, Hansen MK, et al. A controlled study of use of patient-reported outcomes to improve assessment of late effects after treatment for head-and-neck cancer. Radiother Oncol. 2016;119(2):221–8.
- Kroenke K, Talib TL, Stump TE, Kean J, Haggstrom DA, DeChant P, et al. Incorporating PROMIS Symptom Measures into Primary Care Practice-a Randomized Clinical Trial. J Gen Intern Med. 2018;33(8):1245–52.
- Kumar RM, Fergusson DA, Lavallée LT, Cagiannos I, Morash C, Horrigan M, et al. Performance Feedback May Not Improve Radical Prostatectomy Outcomes: The Surgical Report Card (SuRep) Study. J Urol. 2021;206(2):346–53.
- 50. Mallen CD, Nicholl BI, Lewis M, Bartlam B, Green D, Jowett S, et al. The effects of implementing a point-of-care electronic template to prompt routine anxiety and depression screening in patients consulting for osteoarthritis (the Primary Care Osteoarthritis Trial): A cluster randomised trial in primary care. PLoS Med. 2017;14(4): e1002273.
- McCambridge J, Day M. Randomized controlled trial of the effects of completing the Alcohol Use Disorders Identification Test questionnaire on self-reported hazardous drinking. Addiction. 2008;103(2):241–8.
- 52. Reiber GE, Au D, McDonell M, Fihn SD. Diabetes quality improvement in Department of Veterans Affairs Ambulatory Care Clinics: a grouprandomized clinical trial. Diabetes Care. 2004;27(Suppl 2):B61–8.
- Richardson J, Chan D, Risdon K, Giles C, Mulveney S, Cripps D. Does monitoring change in function in community-dwelling older adults alter outcome? A randomized controlled trial Clin Rehabil. 2008;22(12):1061–70.

- Rogers SN, Allmark C, Bekiroglu F, Edwards RT, Fabbroni G, Flavel R, et al. Improving quality of life through the routine use of the patient concerns inventory for head and neck cancer patients: main results of a cluster preference randomised controlled trial. Eur Arch Otorhinolaryngol. 2021;278(9):3435–49.
- 55. Santana MJ, Feeny D, Johnson JA, McAlister FA, Kim D, Weinkauf J, Lien DC. Assessing the use of health-related quality of life measures in the routine clinical care of lung-transplant patients. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2010;19(3):371–9.
- Shyu YI, Liang J, Tseng MY, Li HJ, Wu CC, Cheng HS, et al. Comprehensive and subacute care interventions improve health-related quality of life for older patients after surgery for hip fracture: a randomised controlled trial. Int J Nurs Stud. 2013;50(8):1013–24.
- Varagunam M, Hutchings A, Neuburger J, Black N. Impact on hospital performance of introducing routine patient reported outcome measures in surgery. J Health Serv Res Policy. 2014;19(2):77–84.
- Volkmann ER, FitzGerald JD. Reducing gender disparities in post-total knee arthroplasty expectations through a decision aid. BMC Musculoskelet Disord. 2015;16(1):16.
- Wu S, Ell K, Jin H, Vidyanti I, Chou CP, Lee PJ, et al. Comparative Effectiveness of a Technology-Facilitated Depression Care Management Model in Safety-Net Primary Care Patients With Type 2 Diabetes: 6-Month Outcomes of a Large Clinical Trial. J Med Internet Res. 2018;20(4): e147.
- Zaslansky R, Chapman CR, Baumbach P, Bytyqi A, Castro Lopes JM, Chetty S, et al. Improving perioperative pain management: a preintervention and postintervention study in 7 developing countries. Pain Rep. 2019;4(1): e705.
- Partridge T, Carluke I, Emmerson K, Partington P, Reed M. Improving patient reported outcome measures (PROMs) in total knee replacement by changing implant and preserving the infrapatella fatpad: a quality improvement project. BMJ quality improvement reports. 2016;5(1).
- Baker A, Mitchell EJ, Partlett C, Thomas KS. Evaluating the effect of weekly patient-reported symptom monitoring on trial outcomes: results of the Eczema Monitoring Online randomized controlled trial. Br J Dermatol. 2023;189(2):180–7.
- Balestrieri M, Sisti D, Rocchi M, Rucci P, Simon G, Araya R, de Girolamo G. Effectiveness of clinical decision support systems and telemedicine on outcomes of depression: a cluster randomized trial in general practice. Fam Pract. 2020;37(6):731–7.
- Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol. 2016;34(6):557–65.
- 65. Berdal G, Sand-Svartrud AL, Linge AD, Aasvold AM, Tennebø K, Eppeland SG, et al. Bridging gaps across levels of care in rehabilitation of patients with rheumatic and musculoskeletal diseases: Results from a stepped-wedge cluster randomized controlled trial. Clin Rehabil. 2023;37(9):1153–77.
- 66. Berinstein JA, Cohen-Mekelburg SA, Greenberg GM, Wray D, Berry SK, Saini SD, et al. A Care Coordination Intervention Improves Symptoms But Not Charges in High-Risk Patients With Inflammatory Bowel Disease. Clin Gastroenterol Hepatol. 2022;20(5):1029–38 e9.
- Cooley ME, Mazzola E, Xiong N, Hong F, Lobach DF, Braun IM, et al. Clinical Decision Support for Symptom Management in Lung Cancer Patients: A Group RCT. J Pain Symptom Manage. 2022;63(4):572–80.
- Dhingra L, Schiller R, Teets R, Nosal S, Dieckmann NF, Ginzburg R, et al. Pain Management in Primary Care: A Randomized Controlled Trial of a Computerized Decision Support Tool. Am J Med. 2021;134(12):1546–54.
- Dobscha SK, Corson K, Hickam DH, Perrin NA, Kraemer DF, Gerrity MS. Depression decision support in primary care: a cluster randomized trial. Ann Intern Med. 2006;145(7):477–87.
- Ferrell B, Chung V, Hughes MT, Koczywas M, Azad NS, Ruel NH, et al. A Palliative Care Intervention for Patients on Phase 1 Studies. J Palliat Med. 2021;24(6):846–56.
- Fortmann AL, Walker C, Barger K, Robacker M, Morrisey R, Ortwine K, et al. Care Team Integration in Primary Care Improves One-Year Clinical and Financial Outcomes in Diabetes: A Case for Value-Based Care. Popul Health Manag. 2020;23(6):467–75.

- Garduño-López AL, Acosta Nava VM, Castro Garcés L, Rascón-Martínez DM, Cuellar-Guzmán LF, Flores-Villanueva ME, et al. Towards Better Perioperative Pain Management in Mexico: A Study in a Network of Hospitals Using Quality Improvement Methods from PAIN OUT. J Pain Res. 2021;14:415–30.
- 73. Holm I, Pripp AH, Risberg MA. The Active with OsteoArthritis (AktivA) Physiotherapy Implementation Model: A Patient Education, Supervised Exercise and Self-Management Program for Patients with Mild to Moderate Osteoarthritis of the Knee or Hip Joint. A National Register Study with a Two-Year Follow-Up. J Clin Med. 2020;9(10).
- Howell D, Li M, Sutradhar R, Gu S, Iqbal J, O'Brien MA, et al. Integration of patient-reported outcomes (PROs) for personalized symptom management in "real-world" oncology practices: a population-based cohort comparison study of impact on healthcare utilization. Support Care Cancer. 2020;28(10):4933–42.
- Howell D, Rosberger Z, Mayer C, Faria R, Hamel M, Snider A, et al. Personalized symptom management: a quality improvement collaborative for implementation of patient reported outcomes (PROs) in "real-world" oncology multisite practices. Journal of patient-reported outcomes. 2020;4(1):47.
- 76. Livanainen S, Ravichandra R, Jekunen A, Arokoski R, Mentu S, Lang L, et al. ePRO symptom follow-up of colorectal cancer patients receiving oxaliplatin-based adjuvant chemotherapy is feasible and enhances the quality of patient care: a prospective multicenter study. J Cancer Res Clin Oncol. 2023;149(10):6875–82.
- 77. Jayakumar P, Moore MG, Furlough KA, Uhler LM, Andrawis JP, Koenig KM, et al. Comparison of an Artificial Intelligence-Enabled Patient Decision Aid vs Educational Material on Decision Quality, Shared Decision-Making, Patient Experience, and Functional Outcomes in Adults With Knee Osteoarthritis: A Randomized Clinical Trial. JAMA Netw Open. 2021;4(2): e2037107.
- Ngo E, Truong MB, Wright D, Nordeng H. Impact of a Mobile Application for Tracking Nausea and Vomiting During Pregnancy (NVP) on NVP Symptoms, Quality of Life, and Decisional Conflict Regarding NVP Treatments: MinSafeStart Randomized Controlled Trial. JMIR Mhealth Uhealth. 2022;10(7): e36226.
- Patel MI, Kapphahn K, Dewland M, Aguilar V, Sanchez B, Sisay E, et al. Effect of a Community Health Worker Intervention on Acute Care Use, Advance Care Planning, and Patient-Reported Outcomes Among Adults With Advanced Stages of Cancer: A Randomized Clinical Trial. JAMA Oncol. 2022;8(8):1139–48.
- Pérez JC, Fernández O, Cáceres C, Carrasco ÁE, Moessner M, Bauer S, et al. An Adjunctive Internet-Based Intervention to Enhance Treatment for Depression in Adults: Randomized Controlled Trial. JMIR Ment Health. 2021;8(12): e26814.
- Pooni A, Brar MS, Anpalagan T, Schmocker S, Rashid S, Goldstein R, et al. Home to Stay: A Randomized Controlled Trial Evaluating the Effect of a Postdischarge Mobile App to Reduce 30-Day Readmission Following Elective Colorectal Surgery. Ann Surg. 2023;277(5):e1056–62.
- Price S, Hamann HA, Halaby L, Trejo J, Rogers FC, Weihs K. Collaborative depression care sensitive to the needs of underserved patients with cancer: Feasibility, acceptability and outcomes. J Psychosoc Oncol. 2023:1–23.
- Rollman BL, Anderson AM, Rothenberger SD, Abebe KZ, Ramani R, Muldoon MF, et al. Efficacy of Blended Collaborative Care for Patients With Heart Failure and Comorbid Depression: A Randomized Clinical Trial. JAMA Intern Med. 2021;181(10):1369–80.
- Skovlund PC, Vind Thaysen H, Schmidt H, Alsner J, Hjollund NH, Lomborg K, Nielsen BK. Effect of patient-reported outcomes as a dialoguebased tool in cancer consultations on patient self-management and health-related quality of life: a clinical, controlled trial. Acta Oncol. 2021;60(12):1668–77.
- 85. Steele Gray C, Chau E, Tahsin F, Harvey S, Loganathan M, McKinstry B, et al. Assessing the Implementation and Effectiveness of the Electronic Patient-Reported Outcome Tool for Older Adults With Complex Care Needs: Mixed Methods Study. J Med Internet Res. 2021;23(12): e29071.
- Tirelli F, Xiao R, Brandon TG, Burnham JM, Chang JC, Weiss PF. Determinants of disease activity change over time in Enthesitis related arthritis: effect of structured outcome monitoring and clinical decision support. Pediatr Rheumatol Online J. 2020;18(1):79.

- Velikova G, Keding A, Harley C, Cocks K, Booth L, Smith AB, et al. Patients report improvements in continuity of care when quality of life assessments are used routinely in oncology practice: secondary outcomes of a randomised controlled trial. Eur J Cancer. 2010;46(13):2381–8.
- Wylde V, Bertram W, Sanderson E, Noble S, Howells N, Peters TJ, et al. The STAR care pathway for patients with pain at 3 months after total knee replacement: a multicentre, pragmatic, randomised, controlled trial. Lancet Rheumatol. 2022;4(3):e188–97.
- Barbera L, Sutradhar R, Howell D, Sussman J, Seow H, Dudgeon D, et al. Does routine symptom screening with ESAS decrease ED visits in breast cancer patients undergoing adjuvant chemotherapy? Support Care Cancer. 2015;23(10):3025–32.
- de Wit M, Delemarre-van de Waal HA, Bokma JA, Haasnoot K, Houdijk MC, Gemke RJ, Snoek FJ. Monitoring and discussing health-related quality of life in adolescents with type 1 diabetes improve psychosocial well-being: a randomized controlled trial. Diabetes Care. 2008;31(8):1521–6.
- 91. Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Healthrelated quality-of-life assessments and patient-physician communication: a randomized controlled trial. JAMA. 2002;288(23):3027–34.
- Hilarius DL, Kloeg PH, Gundy CM, Aaronson NK. Use of healthrelated quality-of-life assessments in daily clinical oncology nursing practice: a community hospital-based intervention study. Cancer. 2008;113(3):628–37.
- Rosenbloom SK, Victorson DE, Hahn EA, Peterman AH, Cella D. Assessment is not enough: a randomized controlled trial of the effects of HRQL assessment on quality of life and satisfaction in oncology clinical practice. Psychooncology. 2007;16(12):1069–79.
- Weingarten SR, Kim CS, Stone EG, Kristopaitis RJ, Pelter M, Sandhu M. Can peer-comparison feedback improve patient functional status? Am J Manag Care. 2000;6(1):35–9.
- Buckley L, Ware E, Kreher G, Wiater L, Mehta J, Burnham JM. Outcome Monitoring and Clinical Decision Support in Polyarticular Juvenile Idiopathic Arthritis. J Rheumatol. 2020;47(2):273–81.
- Cross RK, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A Randomized Controlled Trial of TELEmedicine for Patients with Inflammatory Bowel Disease (TELE-IBD). Am J Gastroenterol. 2019;114(3):472–82.
- Davidson KW, Rieckmann N, Clemow L, Schwartz JE, Shimbo D, Medina V, et al. Enhanced depression care for patients with acute coronary syndrome and persistent depressive symptoms: coronary psychosocial evaluation studies randomized controlled trial. Arch Intern Med. 2010;170(7):600–8.
- Frasure-Smith N, Lespérance F, Prince RH, Verrier P, Garber RA, Juneau M, et al. Randomised trial of home-based psychosocial nursing intervention for patients recovering from myocardial infarction. Lancet. 1997;350(9076):473–9.
- 99. Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, et al. A multifaceted intervention to improve treatment of depression in primary care. Arch Gen Psychiatry. 1996;53(10):924–32.
- Katon WJ, Von Korff M, Lin EH, Simon G, Ludman E, Russo J, et al. The Pathways Study: a randomized trial of collaborative care in patients with diabetes and depression. Arch Gen Psychiatry. 2004;61(10):1042–9.
- Katzelnick DJ, Simon GE, Pearson SD, Manning WG, Helstad CP, Henk HJ, et al. Randomized trial of a depression management program in high utilizers of medical care. Arch Fam Med. 2000;9(4):345–51.
- Kronish IM, Moise N, Cheung YK, Clarke GN, Dolor RJ, Duer-Hefele J, et al. Effect of Depression Screening After Acute Coronary Syndromes on Quality of Life: The CODIACS-QoL Randomized Clinical Trial. JAMA Intern Med. 2020;180(1):45–53.
- Maguire R, McCann L, Kotronoulas G, Kearney N, Ream E, Armes J, et al. Real time remote symptom monitoring during chemotherapy for cancer: European multicentre randomised controlled trial (eSMART). BMJ. 2021;374: n1647.
- 104. Patel MI, Ramirez D, Agajanian R, Agajanian H, Bhattacharya J, Bundorf KM. Lay Health Worker-Led Cancer Symptom Screening Intervention and the Effect on Patient-Reported Satisfaction, Health Status, Health Care Use, and Total Costs: Results From a Tri-Part Collaboration. JCO Oncol Pract. 2020;16(1):e19–28.
- 105. Regueiro M, Click B, Anderson A, Shrank W, Kogan J, McAnallen S, Szigethy E. Reduced Unplanned Care and Disease Activity and

Increased Quality of Life After Patient Enrollment in an Inflammatory Bowel Disease Medical Home. Clin Gastroenterol Hepatol. 2018;16(11):1777–85.

- 106. Sharpe M, Walker J, Holm Hansen C, Martin P, Symeonides S, Gourley C, et al. Integrated collaborative care for comorbid major depression in patients with cancer (SMaRT Oncology-2): a multicentre randomised controlled effectiveness trial. Lancet. 2014;384(9948):1099–108.
- Simon GE, Ralston JD, Savarino J, Pabiniak C, Wentzel C, Operskalski BH. Randomized trial of depression follow-up care by online messaging. J Gen Intern Med. 2011;26(7):698–704.
- Unützer J, Katon W, Callahan CM, Williams JW Jr, Hunkeler E, Harpole L, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA. 2002;288(22):2836–45.
- Sandberg T, Conner M. A mere measurement effect for anticipated regret: impacts on cervical screening attendance. Br J Soc Psychol. 2009;48(Pt 2):221–36.
- van Anne-Margreet D, Hester FL, Johan PM, Ewout WS. Random variation and rankability of hospitals using outcome indicators. BMJ Qual Saf. 2011;20(10):869–74.
- 111. Goldstein H, Spiegelhalter DJ. League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. J R Stat Soc Ser A Stat Soc. 2018;159(3):385–409.
- 112. Marshall DA, Jin X, Pittman LB, Smith CJ. The use of patient-reported outcome measures in hip and knee arthroplasty in Alberta. Journal of patient-reported outcomes. 2021;5(2):87.
- 113. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. BMJ Qual Saf. 2014;23(4):290–8.
- 114. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. The Cochrane database of systematic reviews. 2012(6):Cd000259.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.