Patient-Reported Outcomes in Orthopaedics

Article *in* The Journal of Bone and Joint Surgery · March 2018
DOI: 10.2106/jbjs.17.00608

42 authors, including:

Daniel L Christensen
Walter Reed National Military Medical Center
6 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Jonathan Dickens
Walter Reed National Military Medical Center
43 PUBLICATIONS **319** CITATIONS
SEE PROFILE

Brett Owens
Alpert Medical School - Brown University
275 PUBLICATIONS **6,017** CITATIONS
SEE PROFILE

Benjamin Kyle Potter
Walter Reed National Military Medical Center
198 PUBLICATIONS **3,500** CITATIONS
SEE PROFILE

Current Concepts Review

Patient-Reported Outcomes in Orthopaedics

MOTION Group*

- Patient-reported outcomes (PROs) assess a patient's perspective of health, function, and quality of life associated with health conditions and medical interventions.
- Health-care value is the ratio of health outcomes achieved relative to the total cost of care for a medical condition.
- Common PRO designs assess general quality of life, system or region-specific conditions, disease-specific conditions, or mixed outcome measure designs.
- Patient-Reported Outcomes Measurement Information System (PROMIS) measures domains of health (e.g., fatigue, physical function, and depressive symptoms), not disease or injury-specific outcomes, allowing for comparability across conditions and normative scoring.
- PROs offer a unique format to the clinician for understanding the impact of medical conditions or interventions and, as a result, may improve the care provided.

Recent paradigm shifts in the delivery and quality assessment of health care in the United States emphasize patient-centered care and health-care value[1-3]. Patient-reported outcomes (PROs) can be used to quantify patients' perceptions into a value-based model of health-care results[4,5]. The collection of PROs is becoming increasingly common among health-care providers, medical societies, and payers[6]. PRO data collection has been encouraged by the Centers for Medicare & Medicaid Services, which incentivize PRO collection through voluntary quality reporting initiatives[7]. Within orthopaedics, the value of PROs is evidenced by the fact that the American Board of Orthopaedic Surgery is now collecting PROs as part of the Part II Board Certification process[8].

PROs quantify a patient's perceptions of his or her health and/or response to a medical intervention as it affects the patient's health, function, or quality of life[9]. PROs have the potential to inform health-care providers and policy makers with respect to the health-care value that specific medical and/or surgical interventions provide to patients and the impact that these interventions have on patient well-being and function[10]. Understanding patient well-being and function is a prerequisite toward quantifying improvements in patient care. The purpose of this article was to review PROs that are relevant to orthopaedic health-care value, the basic constructs of PROs, the clinical benefit of PRO data, barriers to implementation, and the future direction of PROs in orthopaedics.

*MOTION (Military Orthopaedics Tracking Injuries and Outcomes Network) Group Members: Daniel L. Christensen, MD, Jonathan F. Dickens, MD, Brett Freedman, MD, Timothy Mauntel, PhD, ATC, Brett D. Owens, MD, Benjamin K. Potter, MD, Matthew Provencher, MD, John M. Tokish, MD, Brian R. Waterman, MD, Ivan Antosh, MD, Jaime Bellamy, DO, Kenneth L. Cameron, PhD, MPH, ATC, Karon F. Cook, PhD, Tobin T. Eckel, MD, Josef K. Eichinger, MD, EStephan J. Garcia, MD, Melvin D. Helgeson, MD, Patrick W. Joyner, MD, Daniel G. Kang, MD, Kelly G. Kilcoyne, MD, Chad A. Krueger, MD, Lance E. LeClere, MD, Xinning Li, MD, Kevin D. Martin, DO, Jeremy R. McCallum, MD, Lucas S. McDonald, MD, MPH&TM, Dionisio Ortiz III, MD, Mark P. Pallis, DO, Stephen A. Parada, MD, Jeanne C. Patzkowski, MD, Matthew A. Posner, MD, Jessica C. Rivera, MD, R. Judd Robins, MD, Nan Rothrock, PhD, Matthew R. Schmitz, MD, Mark Slabaugh, MD, Daniel J. Song, MD, Daniel J. Stinner, MD, Scott M. Tintle, MD, Christopher J. Tucker, MD, Robert A. Waltz, MD, and Kevin Wilson, MD

Disclosure: The authors indicated that no external funding was received for any aspect of this work. The **Disclosure of Potential Conflicts of Interest** form is provided with the online version of the article (http://links.lww.com/JBJS/E597).

Disclaimer: The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of the Army, Department of Defense, or the U.S. Government.

Value-Based Health Care

Value in health care is defined as the ratio of outcomes to cost, but the optimal methods to accurately and reliably assess outcomes are debated[1,11]. A plethora of methods exist for defining health-care outcomes, including clinical measurements determined by health-care professionals (e.g., physical examination findings, radiographic studies, operative time, and length of stay) and those reported by patients (e.g., PRO data)[12,13]. Additionally, evaluations of patient experiences, such as satisfaction or process metrics measuring timeliness of care or patient interactions, are often used as proxies for patient outcomes[14]. The patient experience is an important aspect for quality improvement but is not equivalent to quality of health-care delivered or functional outcomes. For example, patient satisfaction is a broader term, independent of outcomes, which may be high if the process of care delivery is perceived as acceptable, even in the setting of a poor health outcome[15].

Clinical outcomes used for assessment of health-care quality and value require stratification that is based on patient-specific factors. Associated disease conditions, expectations, and behavioral and psychometric baseline characteristics may substantially affect outcomes; consequently, assessment of health-care value improves when outcomes are risk-stratified, allowing for more appropriate comparisons[16-18]. Furthermore, assessment of health-care quality, and thus value, may be strongly influenced by who collects the measurement, the type of data collected, the bias inherent in those who respond to outcome data collection, as well as when the outcomes are collected[19,20]. In addition, translation of health-care outcomes into a quality-based and then value-based equation offers opportunity for substantial variability, if outcomes are not standardized.

The denominator of the health-care value equation is the cost of care for the measured intervention or episode of care. Accurate cost assessments are difficult to calculate because of the multifactorial nature of health care, including, but not limited to, anesthesia, facility, surgeon, implant, and rehabilitation costs, all of which vary widely. Bundled payment initiatives seek to improve health-care value by better aligning financial incentives with quality-based outcomes, while also standardizing costs relative to the traditional fee-for-service model that preferentially rewarded quantity over quality[21]. Consequently, PRO data are increasingly connected to reimbursement through legislative actions, including most recently the Medicare Access and CHIP (Children's Health Insurance Program) Reauthorization Act of 2015, which assigns both financial incentives and penalties on the basis of various quality measures[21-23]. While cost is a critical factor in the overall value equation, this review focuses on the numerator—health-care outcomes. As in all fractions, the denominator side of the equation carries equal weight, but the factors that define costs and their relationship to value are wholly different and outside the scope of this review[24].

Outcome Measures

The fundamental design, types, characteristics, and timing of the collection of PROs influence their clinical interpretation and importance (Table I)[25-28]. Wilson and Cleary proposed a classification system for health outcome measures with 5 levels ranging from simple measures to complex ones that incorporate multiple inputs (intrinsic and extrinsic, controllable and noncontrollable)[18]. This classification model demonstrates the difficulty in interpreting biological functions without considering other factors important to the patient, which highlights the necessity for both PROs and clinician-measured outcome measures. Furthermore, the choice of an outcome measure must be weighed against both the patient's complaint and the physician's intervention. Outcome measures must be appropriately selected to coincide with the purpose of the intervention; selection must account for the type of measure with choices including general quality of life, system-specific measures, and/or PROMIS (Patient Reported Outcomes Measurement Information System) measures.

Common Patient Outcome Measure Designs

Mixed Outcome Measure

Mixed outcome measures incorporate data from answered questions and a clinician-completed physical assessment. There are several limitations with current mixed outcome measures: (1) physical assessments are highly dependent on the physician's examination, (2) the patient's willingness to comply and provide satisfactory effort is required, and (3) most mixed outcome measures are not validated[26]. Furthermore, mixed outcome measures focus on unique descriptors of function (e.g., range of motion and pain) that may or may not directly translate to overall well-being and function.

General Health-Related and Overall Quality-of-Life Outcome Measures

General health-related quality of life includes physical, mental, and social factors that influence an individual's health and determine how the individual's health affects his or her ability to perform the activities of daily living (ADLs)[26,29]. Overall quality-of-life measures attempt to quantify a patient's overall well-being, satisfaction, or happiness with life in general[18,26]. Health-related quality-of-life outcome measures are commonly the preferred means for measuring primary outcomes in clinical research as they focus broadly on the impacts a disease and treatment have on overall health instead and are useful in evaluating changes in health status over time. As there is currently no method by which to customize quality-of-life measurements to each individual patient, the use of clinically validated quality-of-life measures with large sets of normative data may be the best alternative[30,31]. Quality-of-life measures can assess whether a treatment restores a patient to a superior, equivalent, or inferior degree of health compared with a meaningful reference group with or without the disease in question. These measures greatly improve the external validity of the data and are essential in health-care economic studies because of the ability to compare the economic benefit of specific interventions. Thus, they have an important utility for both clinicians and health-care administrators. While general measures have unique benefits, a specific limitation is the lack of specificity and responsiveness to system or disease-specific

TABLE I Psychometric Properties of PROs

Psychometric Property	Definition	Clinical Importance
Validity[25]	The degree that the questionnaire accurately measures the content and/or concept of interest.	
Content validity[26,27]	The degree that the content of interest is systematically evaluated within the questionnaire.	Content validity primarily ensures that the questionnaire meets its intended objective (e.g., evaluation and discrimination) and addresses concepts relevant to the population of interest.
Criterion validity[27]	The degree that the questionnaire score(s) relates to a gold standard.	Criterion validity ensures that the questionnaire can effectively serve as a proxy for the gold-standard of measurement for the measurement of interest.
Construct validity[26,27]	The degree that the questionnaire score(s) relates to other measures in ways that are consistent with theoretically derived a priori hypotheses.	Construct validity must be determined when criterion validity is unable to be determined, as is often the case. This occurs in the absence of a true gold standard for a measurement.
Consistency[27]	The degree that items in a questionnaire are correlated (i.e., measure the same concept).	Consistency allows clinicians to determine how reliably patients respond to questions that ask about similar concepts.
Reproducibility[26,27]	The degree that repeated measurements in stable populations yield consistent results.	
Agreement[26,27]	The degree that scores on repeated measures are close to each other (absolute measurement error).	Good agreement (small measurement error) allows clinicians to differentiate between clinically important changes and measurement error.
Reliability[26,27]	The degree that the questionnaire can distinguish patients from each other (relative measurement error).	High reliability allows clinicians to differentiate between patients (e.g., varying degrees of severity of the same condition).
Responsiveness[27]	The degree that a questionnaire can detect clinically important changes over time (i.e., longitudinal validity).	Responsiveness allows clinicians to detect changes that occur over time and differentiate between true changes and measurement error.
Interpretability[26,27]	The degree that a clinician can assign a clinically meaningful qualitative assessment to quantitative score.	Interpretability allows clinicians to interpret the clinical meaningfulness of changes or differences in questionnaire scores (minimally important change [MIC]).
Floor and ceiling effects[26]	The number of respondents who score at the lowest (floor) or highest (ceiling) score on the questionnaire.	Floor and ceiling effects limit the questionnaire's ability to differentiate between patients and to determine change in an individual's health. Selecting population-appropriate questionnaires limits floor and ceiling effects.

deficits in health status and/or interventions, which limits a clinician's ability to assess small differences between patients.

System-Specific Compared with Disease-Specific Outcome Measures

System and disease-specific PRO measures can provide clinicians the ability to assess specific changes in an outcome tied to elements of a body region or a specific pathology, respectively[32,33]. System-specific questionnaires evaluate a specific body system (e.g., hip joint) while disease-specific questionnaires evaluate the impact that a specific diagnosis has on the patient (e.g., hip osteoarthritis). While system-specific questionnaires provide a more precise outcome measure on a single body system, they may lack the granularity needed to identify differences among patients with specific diagnoses. Disease-specific outcome measures generally have better sensitivity than system-specific measures and provide clinicians with the information needed to assess changes in a patient's specific disease[26]. It should be noted that system and disease-specific outcomes generally do not account for the more global psychological and sociological factors in contrast to the general health outcomes that are commonly collected in the orthopaedic setting, such as the Short Form (SF)-36 and SF-12[32].

Patient-Reported Outcomes Measurement Information System (PROMIS)

In 2002, the director of the National Institutes of Health (NIH) initiated the "Roadmap for Medical Research," which led to the development of PROMIS, a tool designed to provide accessible and flexible measures of patient-reported symptoms and health

outcomes covering a range of chronic conditions[34]. PROMIS measures can be used across a multitude of conditions as they are not limited to specific body systems or diseases. They are not focused on overall summaries of well-being, but on specific symptoms and functions, termed *domains,* that are the target of clinical intervention (e.g., depression or physical function). This represents a subtle shift in focus consistent with the World Health Organization's International Classification of Functioning, Disability and Health, focusing health-care delivery away from disease treatment and toward health promotion and function[35].

PROMIS measures are based on item (question) banks that span all levels of the symptom or outcome being measured. Items within an item bank are "calibrated" using item response theory (IRT) models; that is, the characteristics of each item are mathematically defined (e.g., the difficulty of the item is estimated), allowing scoring that is based on probability equations[36,37]. PROMIS uses T-scores with a mean of 50 and a standard deviation of 10, representing the general population of the United States, while subgroup norms for sex and age are also available[38].

Computer adaptive testing (CAT) allows additional flexibility in administration. With CAT administrations, an initial item is presented, and based on the respondent's selection, a computer algorithm selects subsequent items to administer. This tailored administration continues until a stopping rule is reached. The algorithm narrows in on an individual's level, resulting in a reduction in response burden while gaining precision[39-41]. Testing is stopped automatically once a given level of precision (i.e., measurement error) is reached, after a set number of items are administered (a minimum of 4 and a maximum of 12 items), or a combination of the 2.

PROMIS measures allow direct comparison of scores between all CATs with different questions and short forms[36,37]. Furthermore, PROMIS measures can be compared and correlated with scores from legacy measures using PROsetta Stone[42]. Clinical trials utilizing different outcome measures can use PROsetta Stone tables to translate all scores to a common metric, facilitating new comparisons. While PROMIS provides much freedom and flexibility in outcomes assessment and continues to grow in use, there is limited evidence to support widespread adoption or substantial benefit over legacy measures.

Outcomes Selection

The selection and timing of PRO collection is highly dependent on the question being asked and the population being examined[43]. The specificity and depth required of a clinical researcher may be different from that of a private practitioner. Thus, a "one size fits all" approach to all of outcome measurement should be discouraged, although there may be value in defining a minimal data set that minimizes patient and provider burden while optimizing data relevant to the community as a whole. The optimal outcomes platform should be sufficiently broad to allow data comparisons across disease and population samples and sufficiently deep for comparisons within these parameters. The system must have the flexibility necessary to answer the researcher's question. The infrastructure for the collection, storage, and analysis of the selected outcome measures is equally varied, with multiple options available.

Clinical Use and Benefits of PROs

An important feature of any PRO measure is its ability to detect clinically relevant changes over time in the same patient. Quantifying subtle changes allows time for interventions that may change the patient's clinical course. Recently, the International Knee Documentation Committee (IKDC) scale was compared with the PROMIS physical function (PF) CAT at 5 different time points[44]. The IKDC detected changes at only the latter 2 of 4 postoperative visits, while the PROMIS PF CAT detected changes at all postoperative visits. Earlier identification of changes in a patient's perceived health may allow for more effective interventions or reduce the cost and invasiveness of interventions (e.g., getting a patient to therapy sooner for a specific symptom detected by PROs). That study is one in a growing body of literature comparing PROMIS and legacy measures demonstrating positive correlations with decreased question and time burden on patients (see Appendix)[44-51].

Building on previously published work in which legacy foot and ankle instruments were compared with PROMIS measures, Ho et al. identified threshold preoperative PROMIS scores that could predict which patients would achieve or fail to achieve the minimum clinically important difference[52,53]. Similarly, using PROs, Wylie et al. found that, in patients with full-thickness rotator cuff tears, mental health was more highly correlated with shoulder pain and function than tear size, although the postoperative impact of this relationship remains unclear[54]. Utilizing PROs to identify preoperative threshold scores enables surgeons to quantify the severity of patient symptoms and improve indications for surgery. Thus, PROs can provide individual physicians and the orthopaedic community with data that, when combined with relevant clinical measures, improve the ability to risk-stratify patients into those more likely to improve following surgery and those who are less likely to benefit[53]. Thus, PROs may serve as an important adjunct to traditional risk stratification classification methodology (e.g., American Society of Anesthesiologists classification or the American College of Surgeons National Surgical Quality Improvement Program) and provide a more holistic view of a patient's potential for improvement or lack thereof[55].

Efforts to define and improve the quality of health care have precipitated the ongoing transition of PROs from a predominantly research focus toward an increasing utilization in clinical practice[56]. PROs are often documented in research registries that neither connect with the electronic medical record (EMR) nor provide real-time data to the clinician to influence patient care[57]. Thus, most EMRs focus on data from the physician's perspective and lack concrete data from the patient's perspective. In 2007, an ambulatory clinic successfully implemented a PRO collection system utilizing electronic data

entry in the clinic waiting room. A report was generated that contained trend data from prior visits. This report was then available for immediate use during the clinical encounter[58]. The success of this system demonstrated the successful implementation of PROs into busy clinical practices, with reports available in the EMR.

PRO data that are directly populated into the EMR improve patient-physician communication and monitoring of patients' symptoms and trends over time[57-59]. Papuga et al. recently reported on an office-based real-time PRO collection method integrated with the EMR, providing specific examples of PRO data influencing clinical encounters[59]. In addition, Wagner et al. reported on the successful integration of a predominantly home-based PRO collection system into the patient's EMR in an oncology practice[60]. Completion rates for first-time messages were 50%; however, the study demonstrated that completion of PRO collection outside the clinic with results available in the patient's EMR is feasible. Responses that exceeded a predetermined threshold triggered an alert to the patient's care team facilitating the early identification of patients reporting values outside normal expected limits. This earlier identification of suboptimal outcome may reduce resource utilization, save costs, and improve quality of care. These examples offer a variety of methods by which PROs may be integrated into a clinical setting; however, there remains much to be done to identify the optimal method of clinical integration. The collection of PRO data does not strictly equate with improved care; the physician must be able to efficiently utilize these data within clinical practice to facilitate improvements in patient care. The methods by which PRO collection occurs continue to be refined and will likely vary widely on the basis of subspecialty needs, practice environment, as well as patient and physician-specific factors.

Barriers to PRO Utilization

There is considerable potential benefit in utilizing PROs in orthopaedics. However, barriers to PRO clinical implementation exist; thus, identifying and understanding these barriers will aid in successful execution of PRO program development.

The universal implementation of PROs is challenging. Methods for PRO collection and time frames for administration are not standardized or agreed on. Additionally, complete and accurate data collection is limited by existing technological platforms and dedicated manpower to make the collection, tracking, and reporting of PROs possible[61]. Financial and infrastructure requirements for PRO adoption necessitate commitments from individual physicians or health-care organizations. Collaborative outcomes initiatives that aim to provide multisystem or national organizational adoption of PRO processes optimize data collection and efficiency but face unique challenges. These collaborative or national-level programs are costly and may become fragmented from variations in methods, quality, data content, response rates, and follow-up times that may compromise the primary benefit of "big data" collection.

Large-scale aggregated data (i.e., "big data") collection aims to capture common data elements at common time points for similar diseases and treatments to enable meaningful comparisons that can inform evidence-based process improvement with a focus on the comparative effectiveness of interventions. This assumes, however, that "big data" are valid, accurate, complete, and able to overcome multiple biases. The contrast between the lucidity of the concept (i.e., PROs are good and will enable and improve quality of care) and the ambiguity of implementation threatens the impact of such a process. Accordingly, it remains paramount for physicians to both engage as stakeholders in the redesign of health-care delivery and commit to a value-based model.

A major concern is related to the use of paper surveys for PRO data collection. Higher response rates are observed with the use of electronic surveys and staff involvement[62,63]. Risk factors reported for poor paper survey completion in a population of patients managed with arthroplasty include an age of >75 years, Hispanic ethnicity or black race, Medicare insurance, and revision arthroplasty[62]. Paper measures require time for score calculation and making results accessible to the clinician in addition to the added burden of physical storage.

Another concern is the logistics of collecting PRO data from patients. The time required for collection, the potential for survey fatigue or simply an unwillingness to participate, as well as scoring and review by staff are not unimportant and must be considered potential barriers[64]. Administratively, debate has surrounded the issue of who bears the responsibility of the costs associated with these initiatives as the costs can be substantial, especially considering the increased need for information technology and administrative staff support as well as the potential decreased volume from flow limitations[61]. Interestingly, in the now retired Physician Quality Reporting System, Duncan et al. demonstrated that financial penalties associated with not collecting PROs were far less than the costs associated with PRO collection and reporting[61]. PRO collection is being debated as a potential billable expense that can be reimbursed by insurance and government payers[65]. Financial incentives for both providers and patients may substantially improve motivation on both sides of PRO collection and minimize this barrier. An alternative consideration is the use of an independent contractor, separated from the physician or treating hospital, to collect postoperative PROs[65]. The downsides of this model include physicians losing access to valuable patient data in addition to a data set never seen or evaluated by the treating provider, impacting financial reimbursement.

One additional and major challenge is the absence of agreement on the optimal PROs for orthopaedic practice and research. Each subspecialty and anatomic area uses multiple PROs without consensus. Additionally, the minimal clinically important difference is not always clearly defined, which may limit how best to use data even after collection. The promise of PROMIS is that the embracement of a universal PRO can allow for comparisons across disciplines, both within the field of orthopaedic surgery and beyond. Whether PROMIS will gain widespread acceptance as a stand-alone tool to meet all PRO assessment needs remains to be seen as it does not incorporate disease or injury-specific measurement.

Future Directions

The ability to quantify and compare health-care quality is an evolving concept; however, it is likely that patients' perception of care and outcomes will remain an integral aspect in the determination of the quality of care delivered[66]. The Comprehensive Care for Joint Replacement (CJR) program encourages voluntary reporting of PRO data[65]. While the CJR model impacts only arthroplasty surgeons, it is possible that this requirement will expand to the entire orthopaedic community. Appropriate selection and timing of the collection of PROs are critical to effectively quantify the patient perception of orthopaedic interventions. Incentive systems for patients to complete PRO feedback electronically without requiring a return to the treating physician need to evolve. However, while automation improves the data acquisition process and integration into the medical record, effective utilization of that data ultimately requires human involvement and intentional planning to determine how best to interpret and thus utilize the data. It is critical that orthopaedic surgeons proactively engage in this changing paradigm to ensure that true measures of patient care remain the focus of outcomes assessment.

Appendix

eA A table showing studies comparing the use of PROMIS and legacy measures in orthopaedic populations is available with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJS/E598). ■

E-mail address for J.F. Dickens: jon.f.dickens@gmail.com

ORCID iD for J.F. Dickens: 0000-0002-4189-0287

References

1. Porter ME. What is value in health care? N Engl J Med. 2010 Dec 23;363(26):2477-81. Epub 2010 Dec 8.

2. Suk M, Norvell DC, Hanson B, Dettori JR, Helfet D. Evidence-based orthopaedic surgery: what is evidence without the outcomes? J Am Acad Orthop Surg. 2008 Mar;16(3):123-9.

3. Franklin PD, Allison JJ, Ayers DC. Beyond joint implant registries: a patient-centered research consortium for comparative effectiveness in total joint replacement. JAMA. 2012 Sep 26;308(12):1217-8.

4. Swiontkowski MF, Buckwalter JA, Keller RB, Haralson R. The outcomes movement in orthopaedic surgery: where we are and where we should go. J Bone Joint Surg Am. 1999 May;81(5):732-40.

5. Wright RW. Knee injury outcomes measures. J Am Acad Orthop Surg. 2009 Jan;17(1):31-9.

6. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academies Press; 2001.

7. Stulberg J. The Physician Quality Reporting Initiative—a gateway to pay for performance: what every health care professional should know. Qual Manag Health Care. 2008 Jan-Mar;17(1):2-8.

8. The American Board of Orthopaedic Surgery. Certification examinations part II: patient reported outcomes. https://www.abos.org/certification-exams/part-ii/patient-reported-outcomes.aspx. Accessed 2017 Jun 14.

9. Patrick D, Guyatt GH, Acquadro C. Patient-reported outcomes. In: Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011). The Cochrane Collaboration; 2011. www.handbook.cochrane.org. Accessed 2017 Nov 20.

10. Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013 Jan 28;346:f167.

11. Hung M, Zhang W, Chen W, Bounsanga J, Cheng C, Franklin JD, Crum AB, Voss MW, Hon SD. Patient-reported outcomes and total health care expenditure in prediction of patient satisfaction: results from a national study. JMIR Public Health Surveill. 2015 Sep 23;1(2):e13.

12. Pynsent PB. Choosing an outcome measure. J Bone Joint Surg Br. 2001 Aug;83(6):792-4.

13. Bohl DD, Singh K, Grauer JN. Nationwide databases in orthopaedic surgery research. J Am Acad Orthop Surg. 2016 Oct;24(10):673-82.

14. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988 Sep 23-30;260(12):1743-8.

15. Graham B, Green A, James M, Katz J, Swiontkowski M. Measuring patient satisfaction in orthopaedic surgery. J Bone Joint Surg Am. 2015 Jan 7;97(1):80-4.

16. Alemayehu D, Cappelleri JC. Conceptual and analytical considerations toward the use of patient-reported outcomes in personalized medicine. Am Health Drug Benefits. 2012 Jul;5(5):310-7.

17. Chen RC, Clark JA, Talcott JA. Individualizing quality-of-life outcomes reporting: how localized prostate cancer treatments affect patients with different levels of baseline urinary, bowel, and sexual function. J Clin Oncol. 2009 Aug 20;27(24):3916-22. Epub 2009 Jul 20.

18. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA. 1995 Jan 4;273(1):59-65.

19. Lyman S, Hidaka C. Patient-reported outcome measures-what data do we really need? J Arthroplasty. 2016 Jun;31(6):1144-7. Epub 2016 Mar 21.

20. Dobson F, Bennell K, Hinman R, Roos E, Abbott H, Stratford P, Davis A, Buchbinder R, Snyder-Mackler L, Hansen P, Thumboo J, Henrotin Y. OARSI recommended performance-based tests to assess physical function in osteoarthritis of the hip or knee: authors' reply. Osteoarthritis Cartilage. 2013 Oct;21(10):1625-6. Epub 2013 Jul 21.

21. Centers for Medicare & Medicaid Services (CMS), HHS. Medicare program; Merit-Based Incentive Payment System (MIPS) and Alternative Payment Model (APM) incentive under the physician fee schedule, and criteria for physician-focused payment models. Final rule with comment period. Fed Regist. 2016 Nov 4;81(214):77008-831.

22. Federman AD, Keyhani S. Physicians' participation in the Physicians' Quality Reporting Initiative and their perceptions of its impact on quality of care. Health Policy. 2011 Oct;102(2-3):229-34. Epub 2011 May 31.

23. Doherty RB. Goodbye, sustainable growth rate-hello, merit-based incentive payment system. Ann Intern Med. 2015 Jul 21;163(2):138-9.

24. Dijksman LM, Poolman RW, Bhandari M, Goeree R, Tarride JE; International Evidence-Based Orthopaedic Surgery Working Group. Money matters: what to look for in an economic analysis. Acta Orthop. 2008 Feb;79(1):1-11.

25. Creswell JW, Miller DL. Determining validity in qualitative inquiry. Theory Into Practice. 2000;39(3):124-30.

26. Poolman RW, Swiontkowski MF, Fairbank JCT, Schemitsch EH, Sprague S, de Vet HCW. Outcome instruments: rationale for their use. J Bone Joint Surg Am. 2009 May;91(Suppl 3):41-9.

27. Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42. Epub 2006 Aug 24.

28. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, Stein RE. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res. 2002 May;11(3):193-205.

29. Jackowski D, Guyatt G. A guide to health measurement. Clin Orthop Relat Res. 2003 Aug;413:80-9.

30. Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, Rothendler JA, Kazis LE. Updated U.S. population standard for the Veterans RAND 12-Item Health Survey (VR-12). Qual Life Res. 2009 Feb;18(1):43-52. Epub 2008 Dec 3.

31. Ware JE Jr. SF-36 health survey update. Spine (Phila Pa 1976). 2000 Dec 15;25(24):3130-9.

32. Smith MV, Klein SE, Clohisy JC, Baca GR, Brophy RH, Wright RW. Lower extremity-specific measures of disability and outcomes in orthopaedic surgery. J Bone Joint Surg Am. 2012 Mar 7;94(5):468-77.

33. Smith MV, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. J Bone Joint Surg Am. 2012 Feb 1;94(3):277-85.

34. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M; PROMIS Cooperative Group. The Patient-Reported Outcomes

Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007 May;45(5)(Suppl 1): S3-11.

35. Kostanjsek N. Use of The International Classification of Functioning, Disability and Health (ICF) as a conceptual framework and common language for disability statistics and health information systems. BMC Public Health. 2011 May 31;11 (Suppl 4):S3.

36. Brodke DJ, Hung M, Bozic KJ. Item response theory and computerized adaptive testing for orthopaedic outcomes measures. J Am Acad Orthop Surg. 2016 Nov;24 (11):750-4.

37. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol. 2005 Sep-Oct;23(5)(Suppl 39):S53-7.

38. HealthMeasures. Interpret scores: PROMIS. 2017. http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis. Accessed 2017 Jun 21.

39. Choi SW, Reise SP, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res. 2010 Feb;19(1):125-36. Epub 2009 Nov 26.

40. Fries JF, Witter J, Rose M, Cella D, Khanna D, Morgan-DeWitt E. Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. J Rheumatol. 2014 Jan;41(1):153-8. Epub 2013 Nov 15.

41. Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. J Orthop Res. 2011 Jun;29(6):947-53. Epub 2011 Mar 15.

42. PROsetta Stone. What is PROsetta Stone? http://www.prosettastone.org/Pages/default.aspx. Accessed 2017 Aug 6.

43. Zarins B. Are validated questionnaires valid? J Bone Joint Surg Am. 2005 Aug;87(8):1671-2.

44. Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. J Orthop Res. 2014 Jun;32(6):793-801. Epub 2014 Feb 14.

45. Anthony CA, Glass NA, Hancock K, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS instruments in patients with shoulder instability. Am J Sports Med. 2017 Feb;45(2):449-53. Epub 2016 Oct 21.

46. Beckmann JT, Hung M, Bounsanga J, Wylie JD, Granger EK, Tashjian RZ. Psychometric evaluation of the PROMIS physical function computerized adaptive test in comparison to the American Shoulder and Elbow Surgeons score and Simple Shoulder Test in patients with rotator cuff disease. J Shoulder Elbow Surg. 2015 Dec;24(12):1961-7. Epub 2015 Aug 28.

47. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent validity of the Patient-Reported Outcomes Measurement Information System's physical function computerized adaptive test for the knee and shoulder injury sports medicine patient population. Arthroscopy. 2017 Mar;33(3):608-16. Epub 2016 Dec 10.

48. Tyser AR, Beckmann J, Franklin JD, Cheng C, Hon SD, Wang A, Hung M. Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. J Hand Surg Am. 2014 Oct;39(10):2047-2051.e4. Epub 2014 Aug 16.

49. Overbeek CL, Nota SPFT, Jayakumar P, Hageman MG, Ring D. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. Clin Orthop Relat Res. 2015 Jan;473(1):311-7. Epub 2014 Aug 7.

50. Hung M, Baumhauer JF, Brodsky JW, Cheng C, Ellis SJ, Franklin JD, Hon SD, Ishikawa SN, Latt LD, Phisitkul P, Saltzman CL, SooHoo NF, Hunt KJ; Orthopaedic Foot & Ankle Outcomes Research (OFAR) of the American Orthopaedic Foot & Ankle Society (AOFAS). Psychometric comparison of the PROMIS physical function CAT with the FAAM and FFI for measuring patient-reported outcomes. Foot Ankle Int. 2014 Jun;35(6):592-9. Epub 2014 Mar 27.

51. Hung M, Hon SD, Franklin JD, Kendall RW, Lawrence BD, Neese A, Cheng C, Brodke DS. Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. Spine (Phila Pa 1976). 2014 Jan 15;39(2):158-63.

52. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. J Am Acad Orthop Surg. 2016 Nov;24(11):744-9.

53. Ho B, Houck JR, Flemister AS, Ketz J, Oh I, DiGiovanni BF, Baumhauer JF. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. Foot Ankle Int. 2016 Sep;37(9):911-8. Epub 2016 Aug 16.

54. Wylie JD, Suter T, Potter MQ, Granger EK, Tashjian RZ. Mental health has a stronger association with patient-reported shoulder pain and function than tear size in patients with full-thickness rotator cuff tears. J Bone Joint Surg Am. 2016 Feb 17;98(4):251-6.

55. Namba RS, Inacio MCS, Paxton EW. Risk factors associated with deep surgical site infections after primary total knee arthroplasty: an analysis of 56,216 knees. J Bone Joint Surg Am. 2013 May 1;95(9):775-82.

56. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, Dunn J, Eresian Chenok K, Dunbar M, Overgaard S, Garellick G, Lübbeke A; Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries. Patient-reported outcome measures in arthroplasty registries Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries part II. Recommendations for selection, administration, and analysis. Acta Orthop. 2016 Jul;87(Suppl 1):9-23. Epub 2016 May 26.

57. Ayers DC, Bozic KJ. The importance of outcome measurement in orthopaedics. Clin Orthop Relat Res. 2013 Nov;471(11):3409-11. Epub 2013 Aug 10.

58. Ayers DC, Zheng H, Franklin PD. Integrating patient-reported outcomes into orthopaedic clinical practice: proof of concept from FORCE-TJR. Clin Orthop Relat Res. 2013 Nov;471(11):3419-25.

59. Papuga MO, Dasilva C, McIntyre A, Mitten D, Kates S, Baumhauer JF. Large-scale clinical implementation of PROMIS computer adaptive testing with direct incorporation into the electronic medical record. Health Syst (Basingstoke). 2017;64 (2):1-13.

60. Wagner LI, Schink J, Bass M, Patel S, Diaz MV, Rothrock N, Pearman T, Gershon R, Penedo FJ, Rosen S, Cella D. Bringing PROMIS to practice: brief and precise symptom screening in ambulatory cancer care. Cancer. 2015 Mar 15;121 (6):927-34. Epub 2014 Nov 6.

61. Duncan ST, Jacobs CA, Christensen CP, Nunley RM, Macaulay WB. Is Physician Quality Reporting System worth the cost to report to Center for Medicare and Medicaid Services? J Arthroplasty. 2017 Apr;32(4):1055-7. Epub 2016 Nov 17.

62. Schamber EM, Takemoto SK, Chenok KE, Bozic KJ. Barriers to completion of patient reported outcome measures. J Arthroplasty. 2013 Oct;28(9):1449-53. Epub 2013 Jul 25.

63. Tokish JM, Chisholm JN, Bottoni CR, Groth AT, Chen W, Orchowski JR. Implementing an electronic patient-based orthopaedic outcomes system: factors affecting patient participation compliance. Mil Med. 2017 Jan;182(1): e1626-30.

64. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, Hess R, Miller DM, Reeve BB, Santana M. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res. 2012 Oct;21(8):1305-14. Epub 2011 Nov 3.

65. Centers for Medicare & Medicaid Services (CMS), HHS. Medicare program; Comprehensive Care for Joint Replacement payment model for acute care hospitals furnishing lower extremity joint replacement services. Final rule. Fed Regist. 2015 Nov 24;80(226):73273-554.

66. Patel AA. Patient-reported outcome measures: the promise of PROMIS. J Am Acad Orthop Surg. 2016 Nov;24(11):743.