Outcome Measures in Total Joint Arthroplasty: Current Status, Challenges, and Future Directions

MOHAMAD J. HALAWI, MD

Abstract

Total joint arthroplasty (TJA) is the most commonly performed surgical procedure for the treatment of advanced degenerative joint diseases. Numerous outcome measures for TJA have been developed, which can be reported by physicians, patients, or both. Although outcome tools were traditionally centered on morbidity, mortality, and implant survival, the focus has evolved over recent years to joint-specific, disease-specific, activity-specific, general well-being, and quality of life assessments. However, despite the importance of outcome measures in a time of growing demand for TJA and increased government scrutiny fueled by high implant costs, there remains no “gold standard” method to assess the impact of TJA. The aim of this review is to evaluate the currently available literature on outcome measures in joint arthroplasty, highlighting the strengths and limitations of commonly used instruments. Because outcomes are influenced by a multitude of intangible factors, there is an increasing role for assessing patient satisfaction as a simple way to account for the complex aspects of care. Strategies including proper patient selection and establishing realistic expectations preoperatively are critical to optimizing outcomes. In an era of increasing public scrutiny, the current state of conflicting results and variable correlations among outcome measures implores the need for a standardized system that should incorporate a metric for patient satisfaction adjusted for preoperative expectations. [Orthopedics. 2015; 38(8):e685-e689.]
Outcome measures provide a common language among providers to assess the success or failure of surgery. With a growing demand for total joint arthroplasty (TJA) in an era of increasing government scrutiny, outcome measures become imperative tools to assess the cost-effectiveness of TJA, ensuring accountability and providing justification for health care expenditure. Outcome measures are also valuable quality control tools for improving surgical techniques and implant design.

Although early TJA outcome measures were traditionally centered on morbidity, mortality, and implant survival, the focus has evolved over recent years to emphasize functional performance, well-being, and satisfaction due to the advancements in operative techniques and postoperative care pathways. Currently, outcome measures can be reported by physicians, patients, or both. Physician-reported outcomes assess objective findings, such as range of motion, strength, stability, alignment, and ability to perform specific activities. On the other hand, patient-reported outcomes assess more subjective aspects, such as pain level, vitality, self-esteem, and satisfaction.

The discordance between physician- and patient-reported outcomes has been widely reported, with a growing number of studies favoring patient-based measures. Assessment of patients’ perceived health status and quality of life has been shown to be at least as reliable as physician-reported measures but without the high interobserver variability. Patient-reported outcomes also tend to have stronger long-term responsiveness and are less time intensive.

Such measures do not necessarily require clinic visits and hence have the advantage of decreasing the number of patients lost to follow-up. Furthermore, health care audits are now increasingly requiring assessment of patient-reported outcomes, which can be a discriminating factor in commissioning and funding for hospitals, as is the case in the United Kingdom.

**Current Outcome Measures**

Numerous outcome measures for TJA have been described. These can be categorically classified as joint-specific, disease-specific, activity-specific, or general well-being and quality of life tools. Joint-specific outcome tools include the Charnley Hip Score, Oxford Hip and Knee Scores, Harris Hip Score, Hospital for Special Surgery Knee Score, and the Knee Society Score. Although it is ideal to assess the performance of the individual prosthetic joint, such measures—when used alone—often fail to capture the real impact of surgery due to the influence of global health factors.

One example is measuring knee range of motion (ROM), a major component of joint-specific outcome measures. Miner et al showed not only minimal changes in ROM following surgery, but also poor correlation with patient satisfaction during the first postoperative year. In addition, although there was a trend for higher general function with increasing knee ROM, there is no significant correlation between knee ROM and general function.

Disease-specific outcome measures include the Western Ontario and McMaster Universities Arthritis Index (WOMAC), Arthritis Impact Measurement Scale, and AAOS Musculoskeletal Outcomes Data Evaluation and Management System. The WOMAC scoring system is arguably the most widely used in this category. It is a self-administered patient questionnaire specific to hip and knee arthritis consisting of 3 domains: physical function, pain, and stiffness. The validity, reliability, and responsiveness of the WOMAC was originally demonstrated by Bellamy et al in the context of a randomized, controlled clinical trial of 2 nonsteroidal anti-inflammatory drugs in patients with osteoarthritis of the hip and knee. The WOMAC has since undergone several modifications and extensive validation against other scales and is now available in more than 60 languages. Although there is robust evidence about its favorable utility in TJA, the WOMAC has several limitations. It is not specific for a particular joint and can be influenced by factors other than hip and knee osteoarthritis. Licensing fees may be required, and the length of the questionnaire can potentially lead to problems with attrition. There is also only moderate correlation with patient satisfaction and perceived quality of life.

General health status and quality of life measures include the Short-Form health survey 12 and 36 (SF-12 and SF-36), Nottingham Health profile, and EuroQol questionnaire. The SF-36 is arguably the most commonly used in this category. Originally described by Ware and Sherbourne in 1992, the SF-36 is a self-administered patient questionnaire that measures overall well-being based on vitality, physical function, social function, pain, mental health, general health, and limitations due to mental and physical functions. Early reports on the validity and reliability of the SF-36 come from multicenter placebo-controlled clinical trials of arthritis.

Unfortunately, although the SF-36 is useful in measuring changes due to an intervention on a group level, it has substantial floor and ceiling effects, limiting its ability to detect changes on an individual level. The SF-36 is also not detailed enough to assess the functional impact of a surgical intervention. Like the WOMAC, there is only modest correlation with patient satisfaction. Poor questionnaire completion rates, licensing fees, and only moderate correlation with patient satisfaction and perceived quality of life further limit its wide application.

Activity-specific outcome measures include the University of California Los Angeles (UCLA) activity scale, the Activity Rating Scale, and the Tegner score. Among these measures, the UCLA activity scale has been shown to be the most reliable, with high completion rates and no floor effects when using the International Physical Activity Questionnaire as
In the UCLA activity scale, patients rank their activity level from 1 to 10, with 1 defined as absent activity and 10 defined as being able to participate in impact sports. However, when compared with other outcome measures, such as the WOMAC and SF-12, the UCLA activity scale showed low to moderate correlation. In addition, using an accelerometer to objectively measure changes in physical activity after arthroplasty, Harding et al demonstrated no significant changes in activity despite patients reporting improvements in pain and function.

**THE INTANGIBLES**

Several patient variables have gained attention as predictors of outcomes after TJA, including race, ethnicity, gender, and socioeconomic status. Groenwold et al found that among potential candidates for TJA, African American patients had lower expectations of surgical outcomes than white patients. Interestingly, despite substantial improvement after surgery, African American patients continued to have worse outcomes, as measured by SF-36, WOMAC, and quality of well-being (QWB) index.

Similarly, Lavernia et al found that women had lower scores on the SF-36, WOMAC, and QWB index despite reporting greater improvement than men. Barrack et al reported that minority patients and those with low income were more likely to have functional limitations and lower satisfaction rates after TKA. Greene et al reported that patients with a higher education were more likely to have greater quality of life, pain relief, and satisfaction after THA.

Psychological factors, such as depression, anxiety, optimism, social functioning, vitality, and somatization, can also influence the outcomes of TJA. There is increasing evidence that depression and anxiety have a greater impact on postoperative function and pain than do medical comorbidities. Brander et al showed that depression was an independent predictor of worse outcomes at 1 and 5 years after TKA, as measured by the Knee Society score. Similarly, Singh at al found a significant association between anxiety and knee pain after primary and revision TKA at 2 and 5 years postoperatively.

**PATIENT EXPECTATION**

The role of patient expectations in predicting clinical outcomes continues to emerge, with evidence showing that fulfillment of expectations is directly related to patient satisfaction. Patients come to the surgeon with unique needs and expectations that they anticipate to be fulfilled. Interestingly, certain trends in preoperative expectations were observed: higher expectations were reported in patients undergoing TJA at private hospitals, in men with lower body mass indices, and in younger patients living with a spouse or partner.

Other reports have been conflicting, with Mancuso et al showing that older patients, men, and those with worse functional status had higher expectations of surgery. In a recent systematic review of patient expectations after TJA, Haanstra et al found no consistency in the relationship between preoperative expectations and surgical outcomes, in part due to inconsistencies in the definitions of patient expectations.

Still, unrealistic expectations can lead to frustration, which can adversely affect the outcome of an otherwise well-executed TJA. Ghomrawi et al found that one-third of patients had discordantly higher expectations than those of their surgeons. Manning et al found that patients undergoing TJA were overly optimistic and significantly underestimated the time to full recovery. Mancuso et al reported lower satisfaction rates in patients who expected improvement in nonessential activities, such as hiking, dancing, and being off medications. Unrealistic expectations are common and should be moderated by the surgeon during preoperative consultation so that there are clear and realistic goals of surgery to avoid dissatisfaction.

Despite demonstrating the validity and reliability of these patient satisfaction measures, only modest correlation with the WOMAC and SF-36 has been found. This is likely explained by the broad endpoints that patient satisfaction captures. Specifically, although the WOMAC and SF-36 are able to account for pain and physical functioning, they fall short of accounting for meeting of expectations and hospital experience as does patient satisfaction.

The current health care climate of changing patient expectations, rising lawsuits, depreciating insurance reimbursements, and increasing governmental regulations brought on by the Affordable Care Act will likely lead to heightened public scrutiny, imploring the need for outcome measures that incorporate patient satisfaction. Measurement of patient satisfaction not only captures the fulfillment of expectations, but also assesses the quality of health care delivery including patient safety and clinical effectiveness.

**CONCLUSION**

Collins and Roos described attributes of “good” outcome measures to include validity, reliability, accessibility to patients and physicians, minimal administrative burden, relevance to the intended patient group, knowledge of the smallest meaningful score needed to detect change, and responsiveness to change.
of such criteria combined with the complexity of patients’ intangible variables may explain the elusiveness of the gold standard outcome measure, which perhaps may never be developed.

Because patient satisfaction is arguably the most important goal of TJA, it should occupy a central place in the design of future outcome measures. Patient satisfaction is the net outcome and the simplest way to assess the complex multivariable aspects of TJA care. Although hip and knee arthroplasty are regarded as among the most successful and cost-effective operations, success from a patient’s perspective is often dependent on meeting expectations. Therefore, the surgeon should remember that proper patient selection and the establishment of clear and realistic expectations are critical to optimizing outcomes. The conflicting results and variable correlation and available outcome measures in an era of increasing governmental scrutiny and patient expectations implores the need for a standardized system, which should incorporate a metric for patient satisfaction adjusted for preoperative expectations.

REFERENCES


