Measuring Tools for Functional Outcomes in Total Knee Arthroplasty

Robert B. Bourne MD, FRCSC

Published online: 9 September 2008
© The Association of Bone and Joint Surgeons 2008

Abstract Total knee arthroplasty has come under increasing scrutiny attributable to the fact that it is a high-volume, high-cost medical intervention in an era of increasingly scarce medical resources. Health-related quality-of-life outcomes have been developed such that healthcare providers might determine how good an intervention is and whether it is cost-effective. Total knee arthroplasty has been subjected to disease-specific, patient-specific, global health, functional capacity, and cost-to-utility outcome measures. Patient satisfaction is high (90%) after total knee arthroplasty and 93% of patients would have this operative procedure again. Large improvements in preoperative to postoperative WOMAC scores occurred (over 39 of 100 points in 82% of patients). Cost-to-quality outcomes demonstrated total knee arthroplasties are extremely cost-effective. This analysis documents total knee arthroplasty is a highly efficacious procedure that competes favorably with all medical and surgical interventions.

Introduction

We all hope to lead healthy, productive lives with ready access to an effective healthcare system [1]; however, an aging population, new technologies, and competition for resources have led to a worldwide crisis in health care. As such, healthcare interventions have come under increasing scrutiny [12, 14]. Healthcare providers need to know how good an intervention is and whether it is cost-effective [22]. The outcomes movement has evolved to answer many of these questions in a scientifically valid manner [6, 24]. Because TKA is a high-volume, high-cost medical intervention, it is not surprising health-related quality-of-life outcomes have been developed to allow investigators to quantify preoperative to postoperative improvements in TKA patient health status [7, 8, 11, 14, 16]. This research has been valuable to patients, surgeons, and healthcare providers. Patients go to orthopaedic surgeons anticipating the best possible outcomes. For surgeons, outcome studies aid continuous quality assurance, and improve surgical techniques and implant selection. For healthcare providers, outcome studies provide assurance of high-quality, cost-effective care. Unfortunately, however, there are no gold standards in terms of TKA outcome tools.

Ideally, outcome measures used to assess an intervention such as TKA should be valid (measures the proper outcome), reproducible (same value should be obtained on repeated assessments of a stable patient), and responsive to changes in a patient's condition [17, 18].

For assessing TKA, validated outcomes tools include those that are disease-specific (ie, WOMAC, Oxford-12) [3, 11, 12], patient-specific (ie, MACTAR) [27], global health (ie, SF-36, SF-12, Nottingham Health Profile, Sickness Impact Profile, and EuroQol) [8, 20, 25, 28, 29], functional capacity (ie, 6-minute walk, 30-second stair

The author certifies that he has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

The author certifies that his institution has approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

R. B. Bourne (✉)
London Health Sciences Centre, University of Western Ontario,
University Campus, 339 Windermere Road, London, ON,
Canada N6A 5A5
e-mail: robert.bourne@lhsc.on.ca

climb, knee injury and osteoarthritis outcome score [KOOS], pedometry studies, and gait analyses) [8, 10, 15, 23, 30], and cost-to-utility (ie, cost-to-quality adjusted life-years) [21, 22, 26]. For the most part, patient-generated data have been preferred to that which is physician-based. An important addition to TKA outcomes has been the addition of patient satisfaction data [13].

The purpose of this study is to provide a sense of how each of these measures applies to TKA.

Materials and Methods

To illustrate the various validated outcome tools available to assess patients undergoing TKA, we draw from data submitted by surgeons in the province of Ontario and contributed to the Ontario Joint Replacement Registry (OJRR) and the Canadian Joint Replacement Registry (CJRR) [2]. The original intent of this data collection was to allow continuous quality improvement, to help predict the future need for THAs and TKAs, to determine the effect of prolonged waiting times on patient outcomes, and to determine the change score in health-related quality-of-life outcomes that might be necessary for a patient to be satisfied with the procedure. Where examples of knee outcome data were unavailable in the OJRR or CJRR databases, we have relied on published TKA outcomes, either our own or those of others, to provide the necessary information. Much of the data we report has been previously published as noted. All participating patients signed an informed consent for their data to be assessed. The data collection was approved by the Ethics Review Board at the University of Western Ontario.

For patient satisfaction, we used the validated patient satisfaction score developed by Jaeschke et al. [18]. In this satisfaction scale, a patient who was uncertain as to their outcome was classified as 0, whereas those who were dissatisfied were scored from -1 to -7 depending on their degree of dissatisfaction. If they were satisfied, they were graded $+1$ to $+7$ depending on their degree of satisfaction.

For disease-specific TKA outcomes, we most commonly used WOMAC scoring as developed by Bellamy et al. [3] to measure pain, stiffness, and physical function. Other disease-specific outcome measures used elsewhere are discussed, particularly in comparison studies to the WOMAC score (ie, Oxford-12) [12, 13].

For patient-specific scores, the MACTAR outcome tool is often used [27]. The MACTAR outcome tool, originally described for patients with rheumatoid arthritis, provides an individualized functional priority approach for assessing improvement in physical disability in clinical trials. For this outcome, each patient is given prompts such that they can select the top five most disabling physical or social

activities for which they underwent total joint arthroplasty. The measure is scored from a 10-cm visual analog scale with 10 being the most disabling. The instrument therefore can document improvement in level of function of each of these activities postoperatively. We have not applied this approach to TKAs but illustrate the scale for THAs.

Many global health scores are available. Dunbar et al. [13] reported the Nottingham Health Profile and the SF-12 [28] gave the best “test-retest” reliability, but that the SF-36 [29] and the Sickness Impact Profile demonstrated the best internal consistency reliability. In their analysis, the disease site-specific Lequesne, Oxford-12, and WOMAC outcomes also performed well. Recently, we performed a study to determine the effect of patient factors such as gender on TKA outcomes [7], exploring WOMAC and SF-12 mental component scores (MCS) and physical component scores (PCS) in 436 female and 292 male patients in whom 843 TKAs were performed and followed for a mean of 9.5 years (Table 1).

For global health assessment, we commonly used the SF-36 or SF-12 outcomes tool as developed by Ware et al. [28, 29]. More recently, the EuroQol global health outcome tool has been popularized, particularly in Europe, and several authors believe this five-item tool could replace the SF-36 as a result of its higher response rate, accepting its lower responsiveness [9, 13].

A number of tests are available to assess functional outcomes after TKA. For functional capacity, we have commonly used the 6-minute walk [15] and 30-second stair climb [10] to assess patients undergoing TKA. Other functional capacity tools include the KOOS, which is based on the WOMAC score but has been expanded to include the outcomes of pain, activities of daily living, sport and recreation function, and knee-related quality of life [23]. Other functional outcomes of interest include the International Knee Documentation [17], the Lower Extremity Functional Scale [4], and the UCLA activity-level rating [30].

For cost-to-utility data, the cost-to-quality adjusted life-year is often used. This is a difficult test to perform because it requires accurate costing data that can be combined with a perceived change in outcomes [17, 24].

Table 1. Gender comparisons: preoperative to postoperative change scores

Scoring method	Female	Male	p value
WOMAC	26 $\pm$ 23	24 $\pm$ 22	0.28
SF-12 (MCS)	0.49 $\pm$ 12	-1.27 ± 11	0.51
SF-12 (PCS)	7 $\pm$ 11	8 $\pm$ 11	0.12

MCS = mental component score; PCS = physical component score.

Results

Chesworth et al. [11] assessed 3003 patients undergoing total joint arthroplasty whose data were captured in the OJRR and reported the majority of patients (82%) had increases in WOMAC scores of over 39 (of 100 total points) and were satisfied with their procedure (Fig. 1). Overall, 93% would have their TKA again ($n = 2788$), 5% were uncertain ($n = 138$), and 3% would not have the procedure again ($n = 77$). The WOMAC scores document the remarkable improvement afforded by TKA (Figs. 2, 3). WOMAC scores document TKA outcomes are superior if surgery is performed less than 100 days from the decision date (Fig. 4).

The patient-specific MACTAR scale identifies the top five most disabling physical or social activities for which they underwent arthroplasty [8, 27]. Patients undergoing THA are substantially improved by the first postoperative year and retain that improvement over 5 years (Fig. 5) [7].

Fig. 1 These Ontario Joint Replacement Registry data demonstrate the preoperative to postoperative WOMAC change scores with patient satisfaction/transition change scores as outlined by Jaeschke et al. [11] ($n = 3003$).

Fig. 2 Patient severity at decision date as measured by WOMAC scores for TKA ($n = 4437$) is demonstrated [2].

Fig. 3 Patient severity 1 year postoperatively after TKA ($n = 4437$) is shown [2].

Fig. 4 The use of preoperative WOMAC scores to predict postoperative TKA patient outcomes based on the length of time patients waited for their surgery. Note that mild (WOMAC, 60), moderate (WOMAC, 40), and severe (WOMAC, 20) patients all did well if their surgery was performed within 100 days from the decision date [2].

Fig. 5 An example of outcomes using the patient-specific MACTAR patient preference disability questionnaire in patients undergoing THA is shown [8]. When stating the top five reasons for undergoing THA on a 10-centimeter visual analogue scale, the mean preoperative to postoperative substantially improved and remained improved up to four years postoperatively.

Table 2. Results of functional tests performed at 10 years

Condition	Nonresurfaced patella group	Resurfaced patella group	Significance (p value)
Mean number of stairs climbed in 30 seconds	31	20	0.043
Extension torque	63.68	55.67	0.359
Flexion torque	36.92	29.79	0.627
Deep flexion test painful	1/20	2/19	0.517
Patellar glide test painful	1/20	1/19	1.000
Patellar grind test painful	2/20	1/19	1.000
ROM—extension	0	0.7	0.411
ROM—flexion	111.7	114.1	0.491

ROM = range of motion.

We found no differences in pre- to postoperative scores between genders using the validated WOMAC, SF-12 (MCS), and SF-12 (PCS) outcome tools [5].

Using a functional outcome, we reported patients with resurfaced patellae could climb more stairs in 30 seconds than patients with unresurfaced patellae at 10 or more years' followup (Table 2) [10]. The KOOS and UCLA activity-level ratings have also been valuable in these assessments [23, 30].

Cost-to-quality adjusted life-year (QALY) is less for THA and TKA than for other common interventions such as heart transplants and coronary artery bypass (Fig. 6) [8, 19, 21]. Although controversial, some investigators have indicated interventions with a cost-to-QALY of less than \$20,000 are so cost-effective that there should be no

Fig. 6 A comparison of cost-to-utility adjusted life years (cost-to-QALY) of various medical and surgical interventions. Treatments < \$20,000 are considered extremely cost-effective, those between \$20,000 - \$100,000 moderately so and those > \$100,000 expensive. (HT = medical treatment of moderate hypertension, CAB = coronary artery bypass, Hemodial = hemodialysis, L. Transp = liver transplantation, HIV = medical treatment of human immune virus infections). HT = medical treatment of moderate hypertension; CAB = coronary artery bypass; Hemodial = hemodialysis; L. Transp = liver transplantation; HIV = medical treatment of human immune virus infections [8].

question about their use in clinical practice from a cost-benefit perspective [21]. The same researchers suggest those interventions with a cost-to-QALY between \$20,000 and \$100,000 are moderately cost-effective and probably should be used. These researchers have questioned the funding of interventions with a cost-to-QALY of more than \$100,000. Obviously, there are substantial ethical and patient concerns about such an approach.

Discussion

The ability to measure the performance of medical devices such as TKA has led to greater acceptance of this intervention based on solid evidence that demonstrates the efficacy, safety, and cost-effectiveness of this surgical intervention as compared with other medical interventions [5]. The application of outcomes data to TKA has fostered continuous quality improvement not only in surgical techniques, but also in implant selection. Using high-level methodology, researchers have and will be able to plan randomized clinical trials to provide high-quality data on TKA designs and techniques, perform the trial, analyze the results, and change clinical practice based on these outcomes [6, 20].

To date, no single best outcome tool has emerged for TKA studies. There are considerable agreement-validated outcomes tools that are responsive, reliable, and reproducible that should be used [11, 20, 22]. Unfortunately, the widely used Knee Society clinical rating has not been validated. Work is underway to develop a new validated Knee Society scoring system and it is hoped this would be given high priority. As a consequence, the WOMAC and Oxford-12 disease-specific scores are the most frequently used outcomes tools. When comparisons to other medical interventions are needed, the global health outcomes such as the SF-12, SF-36, or EuroQol are useful. Cost-to-utility outcomes are of great interest to healthcare providers, but must be based on very accurate costing data and validated outcome measures. The use of real costing data rather than charges must be used if cost-to-utility studies are to be used and provide meaningful comparative information. Historically, precise costs have been combined with clinical improvement on a validated outcome tool such as time trade-off or standard gamble.

Patient satisfaction data are assuming greater importance in outcome studies. A link between patient expectations and postoperative satisfaction has been established; hence, it seems wise to collect and analyze both preoperative expectation and postoperative satisfaction data (Fig. 1). In many ways, satisfaction data are good proxies for a substantially improved validated disease-specific outcome measure. Using the Jaeschke et al. [18]

scale, patients who rank +5 to +7 usually have TKA change scores in excess of 25 of 100 WOMAC points.

The use of outcome tools in TKA will continue to evolve. There remains a need for the development of a single, patient-generated outcomes tool that combines at least some disease-specific, global health, and functional capacity outcomes. Furthermore, postmarket surveillance of TKA procedures will continue to grow. Already, there are national joint arthroplasty registries in Sweden, Norway, Finland, the United Kingdom, Australia, New Zealand, Canada, and many other countries [2, 5, 13]. This trend should be encouraged to continuously improve the field of TKA. The addition of validated outcomes to these national registries will enhance their already enormous value (Figs. 1–4).

References

1. Advisory Council on Health Info Structure. *Canada Health Infoway: Pass to Better Health*. Health Canada Web site. Available at: <http://www.hc-sc.gc.ca/hcs-sss/pubs/ehealth-esante/1999-paths-voies-fin/index-eng.php>. Accessed December 4, 2007.
2. Annual Report of the Ontario Joint Replacement Registry. *Guideline for Wait Time Thresholds for Total Hip and Knee Replacement Surgery Based on Severity*. Ontario, Canada: Ontario Ministry of Health and Long Term Care; 2005:50.
3. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15:1833–1840.
4. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, preassessment properties and clinical application. *Phys Ther*. 1999;9:371–383.
5. Bourne RB, DeBoer D, Hawker G, Kreder H, Mahomed N, Paterson JM, Warner S, Williams J. Total hip and knee replacement. In: *Access to Health Services in Ontario*. Toronto, Ontario: Institute for Clinical Health Services; 2005:91–117.
6. Bourne RB, Maloney WJ, Wright JG. An AOA critical issue—the outcome of the outcomes movement. *J Bone Joint Surg Am*. 2004;86:633–640.
7. Bourne RB, McCalden RW, MacDonald SJ, Mokete L, Guerin J. Influence of patient factors on TKA outcomes at 5 to 11 years follow-up. *Clin Orthop Relat Res*. 2007;464:27–31.
8. Bourne RB, Rorabeck CH, Laupacis A, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. A randomized clinical trial comparing cemented to cementless total hip replacement in 250 osteoarthritic patients: the impact on health related quality of life and cost effectiveness. *Iowa Orthop J*. 1994;14:108–114.
9. Brazier J, Janes N, Kind P. Testing the validity of the EuroQol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res*. 1993;2:169–180.
10. Burnett RS, Haydon C, Rorabeck CH, Bourne RB. Patella resurfacing versus nonresurfacing in total knee arthroplasty. Results of a randomized controlled clinical trial at a minimum of 10 years followup. *Clin Orthop Relat Res*. 2004;428:12–25.
11. Chesworth B, Mahomed N, Bourne RB, Davis A. Willingness to go through surgery again validated the WOMAC clinically important difference from THR/TKR surgery. *J Clin Epidemiol*. 2008 Feb 14 [Epub ahead of print].
12. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br*. 1998;80:63–69.
13. Dunbar ML, Robertsson J, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty. Results of a survey of 3600 patients from the Swedish Knee Arthroplasty Registry. *Acta Orthop Scand*. 2000;71:262–267.
14. Ellwood PM. Shattuck Lecture. Outcomes management. A technology of patient experience. *N Engl J Med*. 1988;318:1549–1556.
15. Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor DW, Berman LB. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Can Med Assoc J*. 1985;132:919–923.
16. Hawker G, Wright J, Coyte P, Paul J, Dittus R, Croxford R, Katz B, Bombardier C, Heck S, Frelind D. Health-related quality of life after knee replacement. *J Bone Joint Surg Am*. 1998;80:163–173.
17. Irrang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, Richmond JC, Shelborne KD. Development and validation of the International Knee Documentation Committee subjective knee form. *Am J Sports Med*. 2001;29:600–613.
18. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10:407–415.
19. Kaper BP, Bourne RB. Total knee replacement is one of the most beneficial procedures both for patients and society. In: Laskin RS, ed. *Controversies in Total Knee Replacement*. New York, NY: Oxford University Press; 2001:1–20.
20. Kirshner B, Guyatt G. Methodological framework for assessing health care indices. *J Chronic Dis*. 1985;38:27–36.
21. Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. Costs of elective total hip arthroplasty during the first year. Cemented versus noncemented. *J Arthroplasty*. 1994;9:481–487.
22. Laupacis A, Rorabeck RB, Feeny D, Tugwell P, Sim DA. Randomized trials in orthopaedics: why, how and when? *J Bone Joint Surg Am*. 1989;71:535–543.
23. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther*. 1998;28:88–96.
24. Sackett DL, Rosenberg WN, Gray JA, Hynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't. *BMJ*. 1996;312:71–72.
25. Sodermann P, Malchau H, Herberts P. Outcome of total hip replacement. A comparison of different measurement methods. *Clin Orthop Relat Res*. 2001;390:163–172.
26. Torrance GW. Measurement of health stake utilities for economic appraisal. *J Health Econ*. 1986;5:1–30.
27. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR patient preference disability questionnaire—an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol*. 1987;14:446–451.
28. Ware J Jr, Kosinski M, Keller SD. A 12-item Short-Form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220–233.
29. Ware J Jr, Sherbourne CD. The MOS 36-item Short-Form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30:473–483.
30. Zahini CA, Schmalzried TP, Szuszcwicz ES, Amstutz HC. Assessing activity in joint replacement patients. *J Arthroplasty*. 1998;13:890–895.