Commonalities, Differences, and Challenges With Patient-derived Outcome Measurement Tools: Function/Activity Scales

Philip C. Noble PhD, Maureen Dwyer PhD, ATC,
Adam Brekke BA

Published online: 2 July 2013
© The Association of Bone and Joint Surgeons® 2013

Abstract

Background There is a critical need to evaluate the success of orthopaedic treatments through valid outcome measures. Previous attempts to express patient outcomes using a single aggregate score led to scores that were ambiguous, often insensitive to change, and poorly correlated with the patient's assessment of the outcome of surgical procedures.

Where Are We Now? Numerous patient-reported outcome measurement tools have been developed for assessment of patients' level of activity and functional status, especially after joint arthroplasty. However, most tools assume an idealized set of prescribed activities independent of the age, activity level, and lifestyle of each individual. Few instruments are designed to capture the

priorities of individual patients, especially those involved in high-demand sporting and recreational activities.

Where Do We Need to Go? We need valid outcome measures that provide a meaningful, individualized assessment of the functional status of each patient, taking into account the lifestyle and expectation of each individual. This advance in outcome measurement will allow clinicians to individualize treatment and provide patients with an accurate estimate of the outcome of alternative treatments and procedures.

How Do We Get There? Much more comprehensive information is needed to characterize the activities, abilities, and physical aspirations of individual patients. This could form a database for the development of predictive models relating individual characteristics to functional outcomes. Statistical tools are needed to minimize the burden on patients in completing questionnaires to access predictive data and to ensure that all outcome assessments are psychometrically valid.

Each author certifies that he or she, or a member of his or her immediate family, has no funding or 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.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request. This work was performed at the Institute for Orthopedic Research and Education, Houston, TX, USA.

P. C. Noble (✉), M. Dwyer
Institute of Orthopedic Research and Education, Suite 2511,
6550 Fannin Street, Houston, TX 77030, USA
e-mail: pnoble@bcm.edu

P. C. Noble
Baylor College of Medicine, Houston, TX, USA

A. Brekke
University of Texas Health Science Center, San Antonio,
TX, USA

Introduction

During the past three decades, there has been increasing emphasis on the importance of activity in maintaining health and longevity [20, 24, 43]. This has led to a shift in the acceptance of exercise as a desirable component of a healthy lifestyle and the elevation of wellness within popular culture, especially among middle-aged patients. This recent focus on exercise and activity has led to a renaissance in sporting and physical activities in older adults with a concomitant expectation that modern medicine should enable older individuals to function at a level previously seen only in those much younger. At the same

time, the number of individuals younger than 55 years of age who require joint arthroplasties for degenerative diseases of the hip and knee has undergone a meteoric increase with no signs of slowing [21, 22, 41]. This represents a new challenge to physicians because many of these patients undergo joint arthroplasty with an expectation that they will be able to resume demanding physical activities without debilitating symptoms [14, 19, 26, 31]. Moreover, a common observation in published studies is that dissatisfaction with the outcome of hip and knee arthroplasty arises from the presence of recurrent joint symptoms and restoration of function at a level below preoperative expectations [3, 32, 34].

In the early days of joint arthroplasty, these procedures were generally reserved for older and less active patients with pain relief as the primary goal. As prosthetic joints were implanted in younger and more active individuals, the incidence of aseptic loosening and component wear threatened the survivorship of joint arthroplasty. The response of many surgeons was to advise their patients to restrict themselves to “acceptable” activities thereby avoiding activities thought to compromise the durability of the implanted devices [9, 16, 23, 27, 35]. In keeping with these recommendations, instruments developed to assess patient outcome after hip or knee arthroplasty were restricted to activities recommended by surgeons, typically activities of daily living, walking on flat surfaces, and stairclimbing, rather than activities actually performed by the patients themselves [33, 35]. Moreover, patient surveys have confirmed that the individual lifestyles of patients, both before and after TKA, are very diverse and that many different recreational and leisure activities involving knee function are of critical importance to each individual [4, 13]. This suggests that any meaningful assessment of functional outcome must be based on both the level of participation of patients in different activities and the personal importance of participation in each specific activity.

In recent years, numerous patient-reported instruments have been proposed for assessment of function after hip or knee arthroplasty [1, 10, 46]. Many of these tools measure multiple dimensions of patient outcomes, including the presence of symptoms, restoration of motion, return to function, and quality of life. The available instruments differ in terms of the degree to which each dimension is assessed. This leads to the conclusion that different instruments may be measuring different underlying traits in some way related to residual symptoms, functional capacity, and patient satisfaction with the outcome of these procedures [4, 8, 10, 12, 25, 28, 34, 38]. Moreover, it is unlikely that all, or perhaps any, of the instruments accurately portray the true functional level of patients after surgery. In fact, a recent systematic review revealed the current measurement tools failed to adequately address the

constructs of activity, participation, and environment in this population [1].

In view of the questions that have arisen regarding the validity and use of the diverse range of tools available for measuring the function and activities of patients after joint arthroplasty, we first present the current tools available for measurement of functional outcome and then discuss the challenges in capturing the diverse activities of patients in the formulation of meaningful patient-reported outcome measures (PROMs). We then conclude by reflecting on improvements in functional assessment that could arise from the application of new technologies and the potential benefits to clinical decision-making.

Where Are We Now?

Numerous measurement tools have been developed for the assessment of patients’ level of activity and functional status in evaluating the outcome of joint reconstruction surgery [1, 6, 10, 11, 34, 37, 38]. However, no consensus has been reached regarding how best to capture such a varied and dynamic aspect of surgical outcome, in part because of inherent differences in the philosophical basis of the instruments that have been developed and the traits that each attempts to measure. This subject is discussed in more detail within Appendix 1 at the end of this article.

Instruments quantifying patients’ level of activity typically document the types, frequency, and intensity of participation of subjects in different activities placing demands on the hip or knee. The earliest functional outcome measures were provider-generated and are epitomized by the Hospital for Special Surgery Knee Rating Scale [18] and the original Knee Society Clinical Rating System [17]. These instruments were empirically developed before the advent of psychometric testing and were formulated on the assumption that certain activities were universal and so could be used as indicators of capabilities of patients undergoing knee arthroplasty in addition to residual symptoms. Thus, these measures contain items quantifying patients’ walking and stairclimbing capabilities with deductions for reliance on walking aids. Attempts to establish the psychometric validity of these instruments have generally been unsuccessful, hence the development of more recent, patient-generated instruments [10, 11, 34, 36, 38].

One of the earliest patient-generated outcome tools used in orthopedics was the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [5, 6]. This instrument quantifies physical function by asking patients to what extent they have had difficulty in performing 17 different daily physical activities in the past 48 hours. However, the activities that are offered have been chosen on the

assumption that all patients perform each activity and that each should be given equal weighting. A few of the published functional outcome instruments contain items relating to higher level activities in patients after total joint arthroplasty [6, 30, 39]. Two instruments developed as extensions of the WOMAC, the Knee Injury and Osteoarthritis Outcome Score and the Hip Disability and Osteoarthritis Outcome Score, have attempted to accommodate younger and more active patients by including items regarding functional, sports, and recreational activities [30, 39]. Both scores have been configured so that it is possible to calculate a separate functional component score by using a Physical Function Short-Form version [11, 37]; however, because each contains only five to seven items, the ability of these measures to quantify function in terms meaningful to individual patients is very limited.

All traditional functional outcome measures are based on the assumption that participation in an activity is indicative of physical capability and, conversely, that absence of participation means that the subject lacks the necessary ability rather than the interest or desire to participate. It is also universally assumed that all subjects perform a given activity with the same intensity and frequency, which is clearly not correct. These concerns have led to the development of scoring systems rating subjects on the basis of the physical demands of the activities they perform. The first of these, the UCLA rating system [2], groups patients into one of 10 descriptive levels according to the frequency and type of activities that they regularly perform, ranging from “wholly inactive and dependent on others” to “regular participation in impact sports.” These data can be useful in determining a patient’s functional

capabilities as well as the types of demands placed on the implanted component [47].

Contemporary assessment instruments aim to identify and quantify outcomes in terms of the patients’ priorities, including pain relief, physical function, and long-term satisfaction [3, 4, 15, 46]. More recently developed patient-specific health assessments, like the Total Hip/Knee Function Questionnaire (THFQ/TKFQ) [44] or the new Knee Society Knee Scoring System (KSS) [32], focus on evaluating patients’ individual concerns, especially with regard to physical function. The THFQ and the TKFQ identify those physical activities that are important to patients undergoing joint arthroplasty, the frequency at which the patient performs each activity, and the prevalence of symptoms during their participation [44]. They include an inventory of different activities spanning the least demanding (walking and standing) to the most demanding (running, downhill skiing). Data derived from these instruments have confirmed that patients participate in a wide range of activities after recovery from joint arthroplasty (Fig. 1) and that an even larger diversity is present in the activities that patients consider personally important (Fig. 2). The challenge in deriving an accurate estimate of the functional deficit experienced by any individual patient is revealed by additional data that reveal no correlation between the importance that patients undergoing knee arthroplasty ascribe to individual activities and the frequency of their participation (Fig. 3). Thus, an instrument consisting of items derived from activities that are commonly performed will fail to include activities of personal importance to many patients, whereas an

Fig. 1 This bar chart depicts the percentage of patients participating in different activities at a minimum of 1 year post-TKA.

Fig. 2 This bar chart shows the distribution of activities cited as “most important” by patients at a minimum of 1 year post-TKA. These data were derived from the validation trial of the New Knee Society Score [30].

Fig. 3 A scatterplot of 24 activities performed by patients who underwent TKA arranged by the percentage of patients regularly participated in each activity versus the percentage considering it to be personally important. Activities in the upper right-hand quadrant were both frequently performed and considered important. These data were derived from the validation trial of the New Knee Society Score [30].

instrument based on activities of most personal importance will neglect those most frequently performed.

This dilemma posed has been addressed in the New KSS, which was derived from an extended version of the TKFQ. To accommodate essential and optional activities, the functional component of the New KSS consists of four subscales: (1) walking and standing (and use of aids) (five items: 30 points); (2) standard activities (eg, climbing stairs; getting up from a couch) (six items: 30 points); (3) advanced activities (eg, squatting, kneeling, running) (five items: 25 points); and (4) discretionary activities: (three items: 15 points).

An important component of this scoring system is the inclusion of discretionary activities by allowing each

subject to choose the three activities from a list of 17 choices that he or she feels is most important [34, 40]. These activities were selected from a master list of 120 activities encompassing sport, recreation, and exercise, which was completed by patients undergoing hip and knee arthroplasty ranging in age from 45 to 91 years. An analysis of data collected from this cohort showed that if patients undergoing TKA are given a choice of 15 different activities, they will find at least two activities that they consider most important in 90% of cases (Fig. 4). Because these instruments include an array of physical activities, they are more apt to discriminate between patients functioning at the extremes of the population distribution, thereby minimizing floor and ceiling effects (Fig. 5). Moreover, the approach

Fig. 4 This plot shows the cumulative increase in the percentage of patients undergoing TKA whose two, three, and four most important activities can be accommodated by a list of one to 15 different activities derived from the total patient population. These data were derived from the validation trial of the New Knee Society Score [30].

taken by the TKFQ and the New KSS overcomes the limitations of many of the earlier outcome instruments by basing the outcome score on those activities that are actually performed by patients rather than those selected by the authors of the questionnaire. In addition, the inclusion of activities that are personally important to each patient more closely estimates the patients' own evaluation of the impact of an intervention on his or her overall quality of life [34, 40].

Although this approach does successfully accommodate many of the discretionary activities that patients consider important, it does not cover the entire spectrum of activities included in surgeon recommendations [16, 35]. Furthermore, patients are not able to write in activities that are not included on the list. An additional risk in the inclusion of "discretionary activities" is that outcome scores will be affected by each individual's self-motivation and aspirations rather than their physical ability to perform given tasks. The authors strongly believe these factors are critically significant in the level of functioning of patients within their environment, and so are essential to a realistic measurement of each individual's "functional outcome," as opposed to his or her "physical capacity." These differences between outcome instruments, in terms of the traits they attempt to measure, should be carefully recognized before selecting PROMs for assessing outcomes.

Current patient-generated functional scores also share several deficiencies that exist for convenience of implementation and/or statistical analysis but hinder their responsiveness to the real functional status of individuals. Disease-specific PROM tools base their estimate of patients' functional status on participation in a spectrum of activities, which have been shown to

Fig. 5 This scattergram shows estimates of the functional status of a population of patients undergoing TKA at a minimum of 1 year followup using two different instruments: (1) a patient-derived score (TKFQ) based on all activities sampled by the survey instrument; and (2) a surgeon-generated score consisting of the Knee Society Function Score [17] based on the patient's ability in walking and climbing stairs.

be important for the specific population. However, to make the tool applicable to the largest number of patients, the level and intensity of activities assessed generally fall on the lower end of functional demand. As such, a number of instruments fall prey to significant ceiling effects during postoperative analyses of function [10, 36]. Ignoring higher level activities in assessment tools greatly limits their ability to detect a change in function as a result of treatment for a large proportion of patients.

Where Do We Need to Go?

There is presently a disconnect between the actual structure of most instruments intended for measurement of functional outcome and the role that most clinicians wish to ascribe to scores generated by the instruments in practice. Because patients are evaluated one individual at a time, clinicians are seeking valid outcome measures that provide a pertinent, meaningful assessment of the functional status of each patient rather than an entire population. This requires a much more comprehensive approach than has been practiced to date using instruments that cover average patients with typical lifestyles and expectations. Moreover, the instruments developed for measuring outcome must be evaluated using psychometric procedures that focus on the accuracy of the instrument in measuring the abilities and functional needs of individuals [4]. This stands in contrast to previous approaches that have evaluated tools in terms of their performance in characterizing the traits of populations and the effectiveness of treatments rather than the outcome of individuals with specific diseases or conditions.

PROM tools could become an extremely valuable addition to personalized medical care if they accurately captured the activities of each patient before and after an episode of impairment to serve as a basis for formulating a treatment plan. A central component of the shared decision-making process could then be a review of each patient's current function in comparison with their pre-disease condition and with their aspirations and expectations. This framework would enable clinicians to provide a rational assessment of the likelihood that a desired outcome will actually be achieved from a particular course of treatment. Based on the connection between patient satisfaction and pretreatment expectations, such an approach is expected to lead to improved outcomes in terms of the satisfaction and confidence of each patient in their outcome and in the expertise of their treating clinician [3, 4]. Although this idealized conception of the "outcome tool of the future" is appealing, the challenge is to create an assessment tool that is accurate and comprehensive without also being cumbersome and impractical for routine use. A solution to this dilemma is still awaited.

PROM instruments that accurately measure patients' current functional status can be also used for routine followup. This opens up great potential applications, including ongoing and Internet-based surveillance of patients' outcomes without the need for office visits. When used in this way, PROM-based tools could signal complications, and other adverse conditions impacting health and joint function, possibly even earlier than the current method of clinic-based followup care. An additional application might be identification of patients at risk for early prosthetic failure based on the frequency and intensity of participation in activities imposing excessive loads on the prosthetic device. Any of these events, flagged through an automated PROM data collection system, could initiate followup communication between the patient and the clinician suggesting further review and consultation, as necessary.

How Do We Get There?

Future efforts in measuring functional outcomes should focus on the challenge of characterizing the activities, abilities, and physical aspirations of each individual subject [13]. This will require separate comprehensive assessments of function and activity in addition to queries regarding the frequency, duration, and intensity of each subject's participation in each activity. An appropriate first step in developing such an instrument would be to create a comprehensive inventory of activities from a large, representative sample of the patient population. We recommend that each activity be rated by all respondents as

something they consider essential to their lifestyle, and/or an activity that they aspire to do, and/or an activity they did immediately before the onset of their disability. Using this database as a reference, a new instrument could offer each patient the choice of a much larger selection of activities than has been possible in the past.

A major obstacle to the use of more individualized and comprehensive outcome assessment tools is the reduction in patient participation as instruments become longer and more complex. The Holy Grail of outcomes measurement is an instrument consisting of a few simple questions possessing all the properties of psychometric validity of much lengthier tools, while requiring virtually no effort to complete. Although such an instrument is unlikely ever to emerge, the mounting pressures from government and insurance providers for increased accountability in health care are driving providers to adopt shorter instruments to obtain data demonstrating some measure of outcome despite loss of sensitivity and questionable validity. An alternative approach is the use of longer instruments to glean more information from each patient with computer-based imputation of missing data to provide statistical estimates of outcome with corresponding confidence intervals. Data enabling comparison of these approaches are critically important to the future of PROMs in orthopaedic practice.

To reduce the burden imposed by lengthy outcome instruments, adaptive testing methodologies should be adopted to focus instruments on ranges of items appropriate for each individual [42]. This process could be guided by an initial assessment of each patient's basic level of activity (eg, sedentary activities; light domestic duties; light recreation and sport; intense exercise and recreation; contact sports) and then progressively refined as more activities are selected from graded subscales. In addition, information regarding the inherent musculoskeletal demands of each activity could be computed on the basis of biomechanical studies and further background research. Combined with information describing the intensity and frequency of each activity, it would be feasible for clinicians to estimate the loading and energy consumption generated by each patient's activity regimen. This information would be particularly useful in estimating the demand imposed by a patient's lifestyle on both their body and their prosthesis and would allow more precise stratification of patients into activity levels for comparison.

The ideal measurement tool would also include queries of activities a patient had done in the past but can no longer perform as well as activities a patient would like to be able to do after surgery. Although some of these activities may not be possible given physical limitations or the restraints of artificial joints, combining this information with the activities a patient does perform will provide a more encompassing depiction of

how the patient perceives their overall ability. In addition, understanding the functional goals of the patient would be critical in the development of appropriate patient-specific postoperative treatment programs.

A final consideration is the need for adaptation of the traditional methods of statistical analysis to meet the demands of individualized assessment of functional outcomes. The concept that instruments should be comprised of items that are suitable for every patient is completely antithetical to the concept of personalized outcome measurement using discretionary items of specific interest to individual patients. Clearly methods of analysis are needed to validate both the content and scoring of instruments in which the items selected for measurement of outcome vary from subject to subject. Because the responses to many items on existing outcome instruments are highly correlated, this challenge is expected to be surmountable through reference to larger data sets collected in support of new instruments to aid interpretation of each item.

Acknowledgments We thank Ms Sophie Fuller-Lafreniere for her assistance with the preparation of the manuscript.

Appendix 1. Activity and Function; Disability and Impairment

Numerous terms have been used to encompass different, related facets of physical activity in evaluating the mechanical performance of the musculoskeletal system [7, 29, 45]. Ambiguity has arisen in the use of this terminology and thus the definition of the traits measured by instruments designed specifically to measure patient activity and functional status. This uncertainty may be attributed to confusion among the following categories of tasks or activities (Fig. 6):

1. Those that a subject is physically capable of performing in the absence of external obstacles;
2. Those that a subject does actually perform in the course of their daily activity; and
3. Those that a subject wants to perform but cannot because of a physical limitation.

In this context, tools that measure “activity” attempt to document the participation of individuals in pursuits regardless of reason, motivation, ability, or social context. Similarly, “impairments” are changes in body function or structure causing a significant reduction in the normal function of the musculoskeletal system (eg, as seen in loss of strength, motion, balance, etc). Assessments of “activity” and “impairment” are not particularly useful in assessing the musculoskeletal status of an individual or the outcome of orthopaedic treatment in the context of meeting the

Fig. 6 This Venn diagram depicts the different categories of activities performed by individuals segmented on the basis of motivation, ability, and necessity.

expectation of the individual. At best, “activity” measurements will indicate how busy a subject is, how many tasks they completed, and potentially how much energy they expended. However, in the context of “outcome” of a treatment or intervention, these data will not inform the observer as to whether the subject was able to perform the activities that were most important to them or consistent with the demands of their chosen vocation or lifestyle. Nor will we be able to determine whether the subject could have been more active but was limited by external factors, including opportunity and accessibility. Estimates of “impairment” ideally show the extent to which an individual is unable to match the musculoskeletal ability of average individuals in terms of joint motion, strength, or the individual’s balance. However, these assessments are silent on the actual demands of the tasks that the individual wishes to perform, and hence tell us nothing about the degree to which the individual’s actual function is compromised.

One construct that contributes to our understanding of the multidimensional nature of disability is the International Classification of Functioning, Disability and Health (ICF) developed by the World Health Organization [45]. This “biopsychosocial” model of disability incorporates influences from biology, the individual, and society to garner a more complete understanding of function. Human functioning is classified by the ICF as having three levels. These include functioning at the level of the body or body part, functioning at the level of the person as a whole, and functioning at the level of the whole person in a societal context. These functions are broadly labeled Body Functions and Structure, Activity, and Participation, respectively. Thus, activity, being the execution of a task or action by an individual, represents a single construct of an individual’s function.

Disability is recognized as a dysfunction at one or more of these levels, termed Impairments, Activity Limitations,

and Participation Restrictions [45]. Impairments are problems that occur in body structure or function. An example would be the loss of ROM that occurs with inflammation as a result of a torn ligament. Activity limitations are difficulties an individual may have in performing activities. For example, as a result of the instability of the joint because of the ligament injury, the individual cannot ambulate without the use of an aid. Participation Restrictions are problems the individual may have with involvement in life situations. An example is the result of the inability of the individual to ambulate independently, he or she can no longer participate in their sport.

References

- Alviar MJ, Olver J, Brand C, Hale T, Khan F. Do patient-reported outcome measures used in assessing outcomes in rehabilitation after hip and knee arthroplasty capture issues relevant to patients? Results of a systematic review and ICF linking process. *J Rehabil Med*. 2011;43:374–381.
- Amstutz H, Thomas B, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am*. 1984;66:228–241.
- Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. *J Arthroplasty*. 1996;11:831–840.
- Bayley KB, London MR, Grunkemeier GL, Lansky DJ. Measuring the success of treatment in patient terms. *Med Care*. 1995;33:AS226–AS35.
- Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. *Semin Arthritis Rheum*. 1989;18:14–17.
- Bellamy N, Buchanan W, Goldsmith C, Campbell J, Stitt L. 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.
- Brandt E, Pope A, ed. *Enabling America: Assessing the Role of Rehabilitation Science and Engineering*. Washington, DC, USA: National Academy Press; 1997.
- Button G, Pinney SA. Meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system? *Foot Ankle Int*. 2004;25:521–525.
- Clifford P, Mallon W. Sports after total joint replacement. *Clin Sports Med*. 2005;24:175–186.
- Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a ‘good’ measure. *Clin Geriatr Med*. 2012;28:367–394.
- Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, Roos EM, Jordan JM, Maillefert JF, Dougados M, Lohmander LS. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. *Osteoarthritis Cartilage*. 2008;16:551–559.
- Dowrick AS, Gabbe BJ, Williamson OD, Cameron PA. Outcome instruments for the assessment of the upper extremity following trauma: a review. *Injury*. 2005;36:468–476.
- Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award. Functional outcome after total knee replacement varies with patient attributes. *Clin Orthop Relat Res*. 2008;466:2597–2604.
- Ghomrawi HM, Franco Ferrando N, Mandl LA, Do H, Noor N, Gonzalez Della Valle A. How often are patient and surgeon recovery expectations for total joint arthroplasty aligned? Results of a pilot study. *HSS J*. 2011;7:229–234.
- Hawker G, Wright J, Coyte P, Paul J, Dittus R, Croxford R, Katz B, Bombardier C, Heck D, Freund D. Health-related quality of life after knee replacement. *J Bone Joint Surg Am*. 1998;80:163–173.
- Healy W, Iorio R, Lemos M. Athletic activity after joint replacement. *Am J Sports Med*. 2001;29:377–388.
- Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res*. 1989;248:13–14.
- Insall JN, Ranawat CS, Aglietti P, Shine J. A comparison of four models of total knee-replacement prostheses. *J Bone Joint Surg Am*. 1976;58:754–765.
- Jourdan C, Poiraudou S, Descamps S, Nizard R, Hamadouche M, Anract P, Boisgard S, Galvin M, Ravaud P. Comparison of patient and surgeon expectations of total hip arthroplasty. *Plos One*. 2012;7:e30195–e30195.
- King DE, Mainous AG III, Carnemolla M, Everett CJ. Adherence to healthy lifestyle habits in US adults, 1988–2006. *Am J Med*. 2009;122:528–534.
- Kurtz S, Ong K, Chan N, Lau E, Halpern M. Prevalence of primary and revision total hip and knee arthroplasty in the United States from 1990 to 2002. *J Bone Joint Surg Am*. 2005;87:1487–1497.
- Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am*. 2007;89:780–785.
- Kuster M. Exercise recommendations after total joint replacement: a review of the current literature and proposal of scientifically based guidelines. *Sports Med*. 2002;32:433–445.
- Larson EB, Bruce RA. Health benefits of exercise in an aging society. *Arch Intern Med*. 1987;147:353.
- Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum*. 1985;28:542–547.
- Mancuso CA, Jout J, Salvati EA, Sculco TP. Fulfillment of patients’ expectations for total hip arthroplasty. *J Bone Joint Surg Am*. 2009;91:2073–2078.
- McGrory B, Stuart M, Sim F. Participation in sports after hip and knee arthroplasty: review of literature and survey of surgeon preferences. *Mayo Clin Proc*. 1995;70:342–348.
- McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31:247–263.
- Nagi S. Some conceptual issues in disability and rehabilitation. In: Sussman MB, ed. *Sociology and Rehabilitation*. Washington, DC, USA: American Sociological Association; 1965:100–113.
- Nilsdotter A, Lohmander LS, Klassbo M, Roos EM. Hip Disability and osteoarthritis Outcome Score (HOOS)—validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord*. 2003;4:10–18.
- Nilsdotter AK, Toksvig-Larsen S, Roos EM. Knee arthroplasty: are patients’ expectations fulfilled? A prospective study of pain and function in 102 patients with 5-year follow-up. *Acta Orthop*. 2009;80:55–61.
- Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452:35–43.
- Noble PC, Gordon MJ, Weiss JM, Reddix RN, Conditt MA, Mathis KB. Does total knee replacement restore normal knee function? *Clin Orthop Relat Res*. 2005;431:157–165.

34. Noble PC, Scuderi GR, Brekke AC, Sikorskii A, Benjamin JB, Lonner JH, Chadha P, Daylamani DA, Scott WN, Bourne RB. Development of a new Knee Society scoring system. *Clin Orthop Relat Res.* 2012;470:20–32.
35. Papalia R, Del Buono A, Zampogna B, Maffulli N, Denaro V. Sport activity following joint arthroplasty: a systematic review. *Br Med Bull.* 2012;101:81–103.
36. Paulsen A, Pedersen A, Overgaard S, Roos EM. Feasibility of 4 patient-reported outcomes measures in a registry setting—a cross-sectional study of 6,000 patients from the Danish Hip Arthroplasty Registry. *Acta Orthop.* 2012;83:321–327.
37. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, Roos EM, Jordan JM, Maillefert JF, Dougados M, Davis AM. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS)—an OARSI/OMERACT initiative. *Osteoarthritis Cartilage.* 2008;16:542–550.
38. Riddle DL, Stratford PW, Bowman DH. Findings of extensive variation in the types of outcome measures used in hip and knee replacement clinical trials: a systematic review. *Arthritis Rheum.* 2008;59:876–883.
39. Roos EM, Roos H, Lohmander LS, Ekdahl C, Beynnon B. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther.* 1998;28:88–96.
40. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The new Knee Society Knee Scoring System. *Clin Orthop Relat Res.* 2012;470:3–19.
41. Singh JA. Epidemiology of knee and hip arthroplasty: a systematic review. *Open Orthop J.* 2011;5:80–85.
42. Smits N, Zitman FG, Cuijpers P, Hollander-Gijsman ME, Carlier IVE. A proof of principle for using adaptive testing in routine outcome monitoring: the efficiency of the Mood and Anxiety Symptoms Questionnaire-Anhedonic Depression CAT. *BMC Med Res Methodol.* 2012;12:1–10.
43. Stampfe MJ, Hu FB, Manson JE, Rimm EB, Willet WC. Primary prevention of coronary heart disease in women through diet and lifestyle. *N Engl J Med.* 2000;341:16–22.
44. Weiss JM, Noble PC, Conditt MA, Kohl HW, Roberts S, Cook KF, Gordon MJ, Mathis KB. What functional activities are important to patients with knee replacements? *Clin Orthop Relat Res.* 2002;404:172–188.
45. World Health Organization. *International Classification of Functioning, Disability and Health: ICF.* Geneva, Switzerland: World Health Organization; 2001.
46. Wright JG. Evaluating the outcome of treatment. Shouldn't we be asking patients if they are better? *J Clin Epidemiol.* 2000;53:549–553.
47. Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. *J Arthroplasty.* 1998;13:890–895.