3-7-2012

Lower extremity-specific measures of disability and outcomes in orthopaedic surgery

Matthew V. Smith
Washington University School of Medicine in St. Louis

Sandra E. Klein
Washington University School of Medicine in St. Louis

John C. Clohisy
Washington University School of Medicine in St. Louis

Geneva R. Baca
Washington University School of Medicine in St. Louis

Robert H. Brophy
Washington University School of Medicine in St. Louis

See next page for additional authors

Current Concepts Review

Lower Extremity-Specific Measures of Disability and Outcomes in Orthopaedic Surgery

Matthew V. Smith, MD, Sandra E. Klein, MD, John C. Clohisy, MD, Geneva R. Baca, BA, Robert H. Brophy, MD, and Rick W. Wright, MD

Investigation performed at the Department of Orthopaedic Surgery, Washington University School of Medicine, St. Louis, Missouri

- Outcome measures may be simple questions or complex measures that assess multiple interrelated domains affecting treatment outcomes.
- Outcome measures should be relevant to patients, easy to use, reliable, valid, and responsive to clinical changes.
- Joint and disease-specific outcome measures have been developed for the hip, knee, and foot and ankle. Many of these measures would benefit from further research into their validity, reliability, and optimal applicability.
- General health measures and activity level scores should be included in outcome assessments after treatment for orthopaedic conditions.

When outcome measures were reported in the early orthopaedic literature, simple metrics such as *return to work* and *patient satisfaction* were used to judge the benefit of intervention for patients. Although these simple metrics continue to be frequently used and are useful[1–5], more complex scoring systems that assess multiple interrelated domains affecting outcomes, such as pain, activities of daily living, and objective physical examination measures, have been developed. These more complex measures may include physician-based assessments, patient-based assessments, or both. They can be joint or disease-specific or focus on general health[6].

In general, outcome measures should be relevant to patients, easy to use, reliable, valid, and responsive to clinical changes[7]. A reliable outcome measure consistently gives the same results under the same testing conditions. Establishing validity for an outcome measure is a complex, multifactorial task that includes evaluation of several facets, including construct validity (Does it measure what it intends to?), criterion validity (Does it correlate to other valid measures?), and content validity (Does it adequately assess the critical features of the problem?). Therefore, the true validity of an outcome is based on a body of evidence evaluating these many facets of validity rather than on a single definitive test. Furthermore, validating an instrument against another evaluates only criterion validity, which does not assess whether the instrument measures what it intends to or adequately assesses the critical features of the problem.

Responsiveness of an outcome measure is its ability to detect changes in a clinical condition. The minimal clinically important difference of an outcome measure is the smallest change in an outcome score that corresponds to a change in a patient's condition[8,9]. If the minimal clinically important difference for an outcome measure is 15 points and a patient records a 10-point change in that measure after treatment, the patient may not note a true change in his or her actual clinical condition. Lastly, outcome measures may have a so-called ceiling effect (an inability to differentiate relatively good or high outcomes) or a floor effect (an inability to differentiate relatively low or poor outcomes). A floor effect occurs when a

Disclosure: None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

subject or a group of subjects performs poorly on an outcome measure at baseline. Because of this poor performance, the outcome measure may not be able to detect further deterioration in their condition as the subjects already approach the minimum score for that measure at baseline.

Our goal in this review is to outline the common outcome measures used to report lower-extremity outcomes in the orthopaedic literature. This is not a comprehensive or detailed evaluation of each measure but rather an overview of why each measure was developed, the conditions each has been reported to assess, and the minimal clinically important difference if available. In addition, general health (Short Form-36 [SF-36] and SF-12) and general orthopaedic outcome measures (Musculoskeletal Functional Assessment [MFA] and Short Musculoskeletal Functional Assessment [SMFA]) are commonly used in conjunction with specific lower-extremity outcome measures. These are summarized in Table I but are not included in the discussion[10–14]. The lower-extremity outcome measures are summarized in Table II (hip), Table III (knee), and Table IV (foot and ankle).

Measure for Knee and Hip Osteoarthritis

Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC LK 3.0)

The Likert 3.0 version of the WOMAC (WOMAC LK 3.0) is widely used and accepted as a disease-specific instrument for osteoarthritis in the lower extremities and for evaluating clinical outcomes after total hip replacement[15]. This is a self-administered instrument with three subscales: pain (0 to 20 points), stiffness (0 to 8 points), and physical function (0 to 68 points). One total score (0 to 96 points) is also reported. There are forty-one items with Likert scale responses from 0 to 4 points[16,17]. The scores are interpreted on a best-to-worst scale, in which lower values indicate less pain and higher function. The scores can be normalized, with 0 indicating severe symptoms and 100 indicating no symptoms and higher function. Importantly, this normalization reverses the interpretation of the score (a higher score indicates no symptoms and higher function). The WOMAC is sensitive to change and easy to use[18–20]; therefore, it is the most commonly used disease-specific outcome measure for osteoarthritis of the hip and knee[18–20].

The WOMAC has undergone rigorous validation and has been used in more than sixty languages[15,21–28], and it has been validated against the SF-36[29]. Considering its focus on older patients with osteoarthritis, it may not be appropriate for use in a young and active population as it is likely to have ceiling effects in this group. The minimal clinically important difference was reported to be 12% of the baseline score or 6% of the maximum score in a study that examined rehabilitation intervention in osteoarthritis[18]. A minimal clinically important difference of 9 to 12 points (on a scale of 0 to 100) has also been reported[30].

Hip-Specific Outcome Measures

There exists an abundance of hip-specific clinical outcome measures as recently summarized by Suk et al.[31]. Below is a summary of the commonly used outcome measures for the hip.

Harris Hip Score

The Harris hip score was initially introduced in 1969 as a research tool to assess the clinical results of mold arthroplasty for traumatic hip arthritis[32]. This scoring system was formulated to measure the important outcome variables applicable to different hip disorders and treatment techniques. Pain (44 points), functional capacity (47 points), deformity correction (4 points), and hip range of motion (5 points) are incorporated into the maximum 100-point scoring system. A higher score indicates higher function. Pain and functional capacity are heavily weighted. The initial description did not specify whether the questionnaire was clinician or patient-administered. This scoring system is a common outcome tool for total hip arthroplasty and hip joint preservation studies, and has been validated for arthroplasty

TABLE I General Health and General Orthopaedic Outcome Measures*

Scale*	Anatomic Region	Measures	Validated	Responder Burden (*no. of questions*)		Target Population		MCID†
				Clinician	Patient	Age (*yr*)	Disorders	
MFA[10]	General orthopaedic	General health and function	Yes	None	101	≥18	Musculoskeletal, fractures, soft-tissue injuries, repetitive motion injuries, and osteoarthrosis	Unknown
SMFA[12]	General orthopaedic	General health and function	Yes	None	46	≥18	Musculoskeletal disorders	Unknown
SF-36[13]	General health	General health	Yes	None	36	Any	General health	
SF-12[14]	General health	General health	Yes	None	12	Any	General health	

*MFA = Musculoskeletal Functional Assessment, SMFA = Short Musculoskeletal Functional Assessment, and SF = Short Form. †MCID = minimal clinically important difference.

TABLE II Hip Outcome Measures for Patients with Osteoarthritis and Patients with a High Activity Level

Scale*	Anatomic Region	Measures	Validated	Responder Burden (*no. of questions*)		Target Population		MCID†
				Clinician	Patient	Age (*yr*)	Disorders	
Hip outcome measures for patients with osteoarthritis								
WOMAC[15]	Hip and knee	Physical function, pain, and stiffness	Yes	None	41	55–71	Hip and knee osteoarthritis and arthroplasty outcomes	12% of baseline score or 6% of max. score; also 9-12 points on 0-100 scale
HHS[32]	Hip	Pain, function, deformity, and range of motion	Yes	5	8	62–71	Traumatic osteoarthritis and/or disorders	7–9
HOOS[36]	Hip	Pain, symptoms, activity limitations, sports and/or recreation, and hip-related quality of life	Yes	None	40	71.5	Primary hip osteoarthritis for total hip arthroplasty	Unknown
Hip outcome measures for patients with high activity level								
HOS[37,38]	Hip	Physical function	Yes	None	19 for ADL and 9 for sports subscale‡	13–66	Labral tears and hip arthroscopy	6–9
MHHS[35]	Hip	Pain and function	No	None	8	Any	Hip-joint preservation surgery	Unknown
Nonarthritic hip score[40]	Hip	Pain, mechanical symptoms, physical function, and level of activity	Yes	None	20	33	Young adults with nonarthritic hip pain	Unknown
UCLA activity score[41]	Lower extremity	Activity level	Yes	10	10	58.4–65.8	Osteoarthritis	Unknown

*WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index, HHS = Harris hip score, HOOS = Hip Disability and Osteoarthritis Outcome Score, HOS = Hip Outcome Score, MHHS = modified Harris hip score, and UCLA = University of California at Los Angeles. †MCID = minimal clinically important difference. ‡ADL = activities of daily living.

patients against the SF-36, WOMAC, Patient-Specific Index (PASI) hip rating scale, and McMaster-Toronto Arthritis questionnaire (MACTAR)[29,33]. The minimal clinically important difference is 4 points with an 8% change in score[34].

The modified Harris hip score was developed in an attempt to construct an outcome tool for more active patients[35]. The questions on deformity correction (4 points) and the range of motion (5 points) were removed, leaving a total of 91 potential points for pain and function. This score is normalized to 100 by multiplying the raw score by 1.1. This score is most commonly reported for hip-joint preservation surgery. The modified score has not been validated. The minimal clinically important difference is unknown.

Hip Disability and Osteoarthritis Outcome Score (HOOS)

The HOOS was developed to construct an instrument with improved responsiveness compared with the WOMAC LK 3.0[36]. This is a self-administered forty-item questionnaire with a 5-point Likert scale used to assess five subscales (pain, symptoms, activities of daily living, sport and recreation function, and hip-related quality of life). All questions on the WOMAC LK 3.0 are contained within the HOOS, and a WOMAC score can be calculated. The additional HOOS questions on sport and recreation function and hip-related quality of life have been shown to be more responsive, especially in patients younger than sixty-six years. Each subscale is scored separately and normalized to 100 points, with a higher score indicating higher function. This score has been validated relative to the SF-36[36]. The minimal clinically important difference is unknown.

High-Activity Hip Outcome

Measures Hip Outcome Score (HOS)

The HOS was developed as an outcome instrument for patients who underwent hip arthroscopy. This is a self-administered questionnaire with two separately scored subscales: activities of

TABLE III Knee Outcome Measures

Scale*	Anatomic Region	Measures	Validated	Responder Burden		Target Population Disorders*	MCID†
				Clinician *(no. of items)*	Patient *(no. of questions)*		
WOMAC[15]	Hip and knee	Physical function, pain, and stiffness	Yes	None	41	Hip and knee osteoarthritis and arthroplasty outcomes	12% of baseline score or 6% max. score; also 9-12 points on 0-100 scale
KOOS[44,45]	Knee	Pain, symptoms, activities of daily living, sport and recreation, and knee-related quality of life	Yes	None	42	Sports-related knee injuries	8–10
IKDC[51] (Subjective Knee Form)	Knee	Symptoms, function, and sports activity	Yes	None	18	All knee conditions	6.3 at 6 mo, and 16.7 at 12 mo
American Knee Society Score[55]	Knee	Pain, stability, range of motion, and function	Yes	3	2	Knee replacement	Unknown
Lysholm Score[58]	Knee	Pain, instability, locking, squatting, limp, support, swelling, and stair-climbing	Yes	None	8	Knee ligament injuries	Unknown
Cincinnati Knee Rating System[65]	Knee	Symptoms, daily and sports activities, physical examination, stability, radiographs, and functional testing	Yes	None	13	ACL injuries, ACL reconstruction, and articular cartilage injury	14 at 6 mo, and 26 at 12 mo
ACL quality of life[69]	Knee	Physical complaints, work, recreation and sports competition, lifestyle, and social and emotional function	Yes	None	32	ACL injuries	Unknown
Tegner activity scale[59]	Knee	Activity level	No	None	1	All knee conditions	Unknown
Marx activity scale[70]	Knee	Activity level	Yes	None	4	All knee conditions	Unknown

*WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index, KOOS = Knee Injury and Osteoarthritis Outcome Score, IKDC = International Knee Documentation Committee, and ACL = anterior cruciate ligament. †MCID = minimal clinically important difference.

daily living and sports. The activities of daily living subscale contains nineteen items (seventeen scored), and the sports subscale contains nine items pertaining to higher-level activities, such as those required in sports[37,38]. The HOS is reported on a scale from 0 to 68 points or a scale from 0% to 100%, with a higher score indicating better function[38]. In validity testing, the HOS activities of daily living and sports subscales had a high correlation to the SF-36 physical function subscale and physical component summary score but a lower correlation to the mental health subscale and mental component summary score. The minimal clinically important difference is 6 to 9 points[39].

Nonarthritic Hip Score

This self-administered questionnaire was developed for younger patients with higher demands than older patients with arthritic hip disease[40]. There are twenty questions (Likert scale responses from 0 to 4), covering four domains, including pain, mechanical symptoms, physical function, and activity level. The ten questions measuring pain and physical function come directly from the WOMAC LK 3.0. The remaining questions focus on mechanical symptoms and levels of activity. The maximum score of 100 indicates normal hip function. The score was initially studied in a young patient cohort with an average age of thirty-three years, and was validated against the Harris hip score and the SF-12[40]. This outcome tool is designed for patients undergoing nonarthroplasty hip surgery. The minimal clinically important difference is unknown.

UCLA Activity Score

The UCLA (University of California at Los Angeles) activity score was introduced as an activity measure for patients undergoing total hip replacement and surface replacement arthroplasty[41]. This instrument evaluates patient activity with ten

TABLE IV Foot and Ankle Outcome Measures

Scale*	Anatomic Region	Measures	Validated	Responder Burden		Target Population Disorders	MCID†
				Clinician *(no. of items)*	Patient *(no. of questions)*		
AOFAS Ankle-Hindfoot scale[92]	Ankle and hindfoot	Pain and function (change after intervention)	No	5	4	Ankle and hindfoot conditions	Unknown
AOFAS Midfoot scale[92]	Midfoot	Pain and function (change after intervention)	No	2	5	Midfoot conditions	Unknown
AOFAS hallux and lesser toe MTP and IP scale[92]	Hallux and lesser toes	Pain and function (change after intervention)	No	5	3	Forefoot conditions	Unknown
FAAM[93]	Foot and ankle	Patient-reported function	Yes	None	29	All foot and ankle	8 on ADL, and 9 on sports subscales
FFI[73]	Foot	Pain and disability	Yes	None	23	Foot diagnoses in rheumatoid patients	Unknown
AOS[88]	Ankle	Pain and disability	Yes	None	18	Ankle arthritis	Unknown
Mazur Ankle Score[111]	Ankle	Patient satisfaction	Yes (construct only)	12	None	Ankle arthrodesis patients	Unknown
VISA-A[85]	Achilles tendon	Pain and function (daily living and sports activities)	Yes	None	8	Achilles tendinopathy	Unknown

*AOFAS = American Orthopaedic Foot & Ankle Society, MTP = metatarsophalangeal, IP = interphalangeal, FAAM = Foot and Ankle Ability Measures, FFI = Foot Function Index, AOS = Ankle Osteoarthritis Score, and VISA-A = Victorian Institute of Sports Assessment-Achilles. †MCID = minimal clinically important difference, and ADL = activities of daily living.

descriptive activity levels (scored on a scale of 1 to 10). Activity levels range from "wholly inactive" (1 point) to "regularly participates in impact sports" (10 points). It has a low clinician and patient burden, and is commonly used to analyze patient activity level and return to sports after joint replacement and joint preservation surgeries. The initial description was reported in a study on patients who were fifty-eight to sixty-six years old and did not specify whether the question was clinician or patient-administered. The activity score has been validated against quantitative assessment of walking activity with a pedometer[42]. The minimal clinically important difference is unknown.

Knee-Specific Outcome Measures

Knee Injury and Osteoarthritis Outcome Score (KOOS)

On the basis of the perceived limitations of the use of the WOMAC in young and active patients, the KOOS was created as an extension of the WOMAC[43]. The KOOS is a patient-based assessment graded on a 5-point Likert scale and is used to evaluate outcomes after sports injuries[44,45]. It evaluates pain, symptoms, activities of daily living, sport and recreation function, and knee-related quality of life[45]. Like the WOMAC, each subscale is summed to give a total score. The total score is transformed to a scale from 0 to 100, with higher scores indicating better outcome. All WOMAC questions are contained within the KOOS so a WOMAC score can be calculated[45]. The KOOS has been used for the assessment of a variety of diagnoses[46-50]. The most sensitive subscales include the pain, sport and recreation, and knee-related quality-of-life domains that have the largest effect size in younger, more active patients[26,46,49]. Although the minimal clinically important difference has not been formally assessed, it has been estimated from the WOMAC that a change of 8 to 10 points in a KOOS subscale is considered clinically important.

International Knee Documentation Committee (IKDC) Subjective Knee Form

The IKDC was established in 1987, and the IKDC form was first published in 1993 as a standardized method to evaluate knee injuries and treatment. In 1997, the Board of the American Orthopaedic Society for Sports Medicine revised the form to become a knee-specific assessment tool rather than a disease and/or condition-specific tool[51]. The result was an eighteen-question patient-rated form that evaluates symptoms, function, and sports activity. The raw scores are summed and transformed to a scale from 0 to 100, with higher scores representing better outcomes[51]. The IKDC subjective form has been validated and shown to be reliable and responsive for a wide range of knee conditions[52,53]. The strength of the IKDC form is that it can be used as a single form to assess any condition involving the knee and thus allow comparison between groups with different diagnoses. In a study including a wide variety of knee conditions, a minimal clinically important difference of 11.5 points on the 100-point scale was sensitive to change in a condition[53]. The minimal clinically important difference of the IKDC subjective form after treatment of

articular cartilage injury is 6.3 at six months and 16.7 at twelve months[54].

American Knee Society Score

The American Knee Society Score was developed in 1989 to assess the results of knee replacement and consists of two parts: (1) a knee score that evaluates pain, stability, and range of motion of the knee (three items) and (2) a function score that mainly assesses walking distance and stair-climbing (two items)[55]. Each portion is scored from 1 to 100, with higher scores indicating better outcome. It has been criticized as a surgeon-assessed score that introduces potential bias. It has been validated and noted to be responsive[56,57]. Although it is used traditionally for knee arthroplasty, it has also been used to assess patients with osteoarthritis who are not undergoing knee replacement. The minimal clinically important difference is unknown.

Lysholm

The Lysholm score was published in 1982 and modified in 1985 to evaluate outcomes after knee ligament surgery with an emphasis on the assessment of instability symptoms[58]. The Lysholm score is a questionnaire consisting of eight items: (1) pain, (2) instability, (3) locking, (4) squatting, (5) limp, (6) support, (7) swelling, and (8) stair-climbing. Each item is assigned a maximum numerical value, with pain and instability carrying the highest weight. The score is summed to give a number between 0 and 100 points[58,59]. The scores are arbitrarily categorized as excellent (95 to 100), good (84 to 94), fair (65 to 83), and poor (<65)[60]. While the Lysholm has been widely used to measure outcomes for knee ligament surgery[60], it has received criticism that it functions better for patients after an anterior cruciate ligament (ACL) reconstruction than for those with other knee conditions[61,62]. Its validity, sensitivity, and reliability have been called into question[4,61,62]. In addition, the Lysholm scoring system may have a ceiling effect since scores tend to be higher compared with other knee outcome measures. The Lysholm score continues to have value, especially for comparison with older literature, given the wide use of the score. It may be of greater value when used in conjunction with other scores and with an activity scale, which may adjust for the ceiling effect. The minimal clinically important difference is unknown.

Cincinnati Knee Rating System

In 1983, the Cincinnati Knee Rating System was described to assess subjective symptoms and activity level[63,64]. Since its original description, it was modified to evaluate thirteen scales and then refined to include six subscales: (1) symptoms (20 points), (2) daily and sports activities (15 points), (3) physical examination findings (25 points), (4) stability (20 points), (5) radiographic findings (10 points), and (6) functional testing (10 points)[65]. The measure is scored on a 100-point scale, with higher scores indicating better outcomes. It has been validated to assess outcomes of ACL injury and reconstruction[4,65]. It is comprehensive and has undergone rigorous assessment[66,67]. The Cincinnati Knee Rating System scores may be lower compared with other outcome measures that assess the same condition[66,67]. This rating system is responsive to clinical change[68]. The minimal clinically important difference of the Cincinnati Knee Rating System after treatment of articular cartilage injury is 14 at six months and 26 at twelve months[54].

ACL Quality of Life

In 1998, this score was established as a disease-specific measure of chronic ACL deficiency[69]. There are thirty-two items that evaluate, with use of a visual analog scale, physical complaints, work-related concerns, recreational activities and sport participation or competition, lifestyle, and social and emotional function. The total score is transformed to a scale from 0 to 100 points with each item weighted equally, with higher scores indicating a better outcome. It has been shown to be valid, reliable, and responsive to clinical change[69]. The minimal clinically important difference is unknown.

Tegner Activity Scale

First published in 1985, the Tegner activity scale was designed to give an objective numerical value to assess a patient's activity level[59]. On a scale from 0 to 10, a score of 0 represents disability secondary to knee problems. Scores between 1 and 5 represent activity levels consistent with work or recreational sports, ranging from sedentary jobs through heavy manual labor. Scores of >5 represent higher-level recreational and competitive sports. A score of 10 was assigned to national or international-level soccer. This scale relates activity to specific sports rather than to the specific skills required to participate in those sports. Thus, it is difficult to apply to all patients across different sports. Formal validation has not been performed. The minimal clinically important difference is unknown.

Marx Activity Scale

The Marx activity scale was developed to create a short, patient-based activity assessment that uses questions designed to assess specific functional activities[70]. It evaluates the patient's level of activity in terms of running, cutting, decelerating, and pivoting. Designed to assess the patient's highest peak activity over the last year, it consists of four questions that are scored from 0 to 4 on the basis of how often the activity is performed. Validated during its development, the scale is easy to use with minimal responder burden. The minimal clinically important difference is unknown.

Foot and Ankle-Specific Outcome Measures

Over sixty outcome and clinical rating systems have been described for conditions of the foot and ankle[71]. Many instruments are disease-specific and useful only in evaluating outcomes after treatment of conditions such as rheumatoid arthritis[72–76], ankle instability[77–83], Achilles tendon disorders[84,85], arthritis[86–88], and calcaneal fractures[89,90]. Below are the instruments specific to the foot and ankle region that are more widely applicable and more frequently used in outcomes research[73,91-93].

American Orthopaedic Foot & Ankle Society (AOFAS) Scales

The AOFAS rating system was initially developed by the AOFAS to report clinical status for any foot or ankle disorder. There are

four separate instruments specific to regions of the foot and ankle: the ankle-hindfoot, midfoot, hallux metatarsophalangeal-interphalangeal, and lesser metatarsophalangeal-interphalangeal scales[92]. The scales contain both clinician-based (range of motion, alignment, gait, and stability) as well as patient-based items (pain, function, walking distance and surfaces, and shoe wear) in three subscales with a maximum score of 100 points, with a lower score corresponding to greater disability. The AOFAS scales have shown low levels of validity when evaluated against SF-36, QALY (quality-adjusted life-year) scores, or the Foot Function Index (FFI)[94–97]. Despite their limitations, they remain some of the most commonly used outcome instruments for the foot and ankle[98,99]. Baumhauer et al. found the hallux and lesser toe instruments reliable in a group of patients with rheumatoid arthritis but did not establish validity of the scales[100]. The AOFAS scales have been established as responsive to change after intervention for foot and ankle conditions and are commonly used for this purpose[101]. The minimal clinically important difference is unknown.

Foot and Ankle Ability Measure (FAAM)

The FAAM was developed to address the need for a universal, validated instrument for the foot and ankle[93]. It is a patient-reported assessment used to evaluate outcomes and health status in patients with foot or ankle disorders[93]. The two subscales, activities of daily living and sports, are included with twenty-nine items, which are transformed to a score of 0 (greatest disability) to 100 (least disability). Validation studies have been performed for a general population of patients with foot and ankle disorders, a group of patients with diabetes mellitus, and athletes with chronic ankle instability[93,102,103]. This instrument is the most extensively validated foot and ankle outcome instrument available, with a minimal clinically important difference of 8 on the activities of daily living subscale and 9 on the sports subscale[93]. The score is sensitive to overall health status and comorbidities. However, limitations may exist in its use with higher-functioning patients because of a potential for a ceiling effect.

Foot Function Index (FFI)

The FFI is a patient-reported assessment tool to measure the impact of pain, disability, and activity restriction related to foot and ankle disorders in patients with rheumatoid arthritis[73]. The initial validation study was performed in a group of patients with rheumatoid arthritis, and further reliability testing has been performed in a general population of patients with a foot disorder[73,76,104,105]. The measure contains twenty-three items in three subscales that are scored to a maximum of 100 points. The items are scored on a visual analog scale, and a higher score indicates greater disability. This measure is a useful index for outcomes in patients with rheumatoid arthritis. Efforts to expand the use of the instrument have resulted in several modified versions (the FFI-5pt[75], FFI Revised, and FFI short form)[106]. While the FFI-5pt (the FFI with verbal rating scales) correlated well with the original FFI, the modified versions have not been used extensively. The minimal clinically important difference is unknown.

Ankle Osteoarthritis Score (AOS)

The AOS is a patient-reported, disease-specific assessment that was developed from a modification of the FFI and is used to measure pain and disability related to osteoarthritis of the ankle[88,107,108]. It has also been used as an outcome measure after treatment of tibial plafond fractures[109,110]. The score is composed of two subscales (pain and disability) with eighteen items scored on a 10-cm visual analog scale, with a maximum score of 100 points. A higher score indicates greater disability. Validation studies indicate high levels of test-retest reliability and criterion and construct validity compared with the SF-36 and WOMAC[88]. While this remains the only validated disease-specific instrument for ankle osteoarthritis, the primary limitation of the score is its limited applicability to a large number of foot and ankle disorders. The minimal clinically important difference is unknown.

Mazur Ankle Score

The Mazur ankle score is a clinician-based outcome instrument derived from the Harris hip-scoring system that is used to evaluate outcome after ankle arthrodesis[111]. Pain, function, and range-of-motion subscales are rated with a twelve-item instrument and a total possible score of 100 points. Results correlate well with successful ankle fusion on radiographs, with higher scores indicating a better outcome. Scores of >60 are associated with patient satisfaction[112]. This instrument has been primarily used in the evaluation of outcome after ankle arthrodesis and tibial plafond fractures[113,114]. It is limited in use because of the disease-specific development of the tool. The minimal clinically important difference is unknown.

Victorian Institute of Sport Assessment-Achilles Questionnaire (VISA-A)

The VISA-A is a patient-reported instrument that measures the clinical severity of Achilles tendinopathy[85,115]. The tool is composed of three subscales with eight items rating pain, activity, and functional status. The maximum score is 100 points, with a lower score indicating greater disability. The tool is useful in rating the clinical severity of Achilles tendinopathy for use in medical decision-making and has been validated against the grade of severity described by Percy and Conochie[115], the grade of severity described by Curwin and Stanish[116], and the Achilles tendinopathy range of severity[85,117]. It has not been validated in the measurement of clinical outcome after surgery. This instrument is also limited in its applicability to other disorders of the Achilles tendon. The Achilles tendon total rupture score was developed on the basis of the VISA-A to assess outcome after Achilles tendon rupture[84]. The minimal clinically important difference is unknown.

Discussion

Orthopaedic surgery has lagged behind other medical and surgical specialties in developing and using outcome measures to study factors that influence the outcomes after treatment[118]. In general, outcome measures for the lower extremity can be useful, but they vary widely in terms of what is known about their validity, reliability, minimal clinically important difference, and applicability

to specific patient populations. It is important to understand how and why these instruments were developed, what they are validated to assess, and how these measures respond to clinical change to avoid misinterpreting reports of patient outcomes.

Investigators performing clinical research should choose measures that have been validated for the disease and/or joint in question. Investigators should also attempt to use instruments that have a reported minimal clinically important difference for ease of interpreting important changes in scores. Furthermore, when choosing outcome measures for clinical research, there is rarely a single most appropriate rating system. Studies should include a measure of general health and an activity scale since rating of disability and outcomes is often affected by psychological and sociological factors that are not accounted for in joint-specific measures. This provides better characterization of patient populations and screens for differences that may influence outcomes. For example, outcomes after an operative intervention in a sedentary patient with multiple medical comorbidities may be affected by health status compared with the same surgical procedure in an active and healthy person. Rigorous outcomes assessment requires a combination of general health measures, activity scales, and condition-specific measures. Lastly, investigators should strive to balance the utility of collecting and analyzing data from multiple outcome instruments with the burden to responders when completing the forms.

In summary, there are a number of clinical outcome measures available for the lower extremity, many of which could benefit from further research into their validity, reliability, and optimal applicability. Investigators should use outcome measures that are valid, reliable, and responsive for assessing the condition being studied. In addition, reviewers and readers should critically evaluate the measures that are used in clinical studies. By increasing their familiarity with these instruments, orthopaedic surgeons are better equipped to design studies of lower-extremity disorders, interpret the data appropriately, and implement the findings into their practices on the basis of sound and informed judgment. ■

Matthew V. Smith, MD
Sandra E. Klein, MD
Robert H. Brophy, MD
Department of Orthopaedic Surgery,
Washington University School of Medicine,
14532 South Outer Forty Drive, Chesterfield, MO 63017.
E-mail address for M.V. Smith: smithmv@wudosis.wustl.edu

John C. Clohisy, MD
Rick W. Wright, MD
Department of Orthopaedic Surgery,
Washington University School of Medicine,
1 Barnes Hospital Plaza, Suite 11300 West Pavilion,
St. Louis, MO 63110

Geneva R. Baca, BA
Department of Orthopaedic Surgery,
Washington University School of Medicine,
660 South Euclid, Campus Box 8233, St. Louis, MO 63110

References

1. Eastlack ME, Axe MJ, Snyder-Mackler L. Laxity, instability, and functional outcome after ACL injury: copers versus noncopers. Med Sci Sports Exerc. 1999;31:210-5.

2. Heckman JD. Are validated questionnaires valid? J Bone Joint Surg Am. 2006;88:446.

3. Neeb TB, Aufdemkampe G, Wagener JH, Mastenbroek L. Assessing anterior cruciate ligament injuries: the association and differential value of questionnaires, clinical tests, and functional tests. J Orthop Sports Phys Ther. 1997;26:324-31.

4. Risberg MA, Holm I, Steen H, Beynnon BD. Sensitivity to changes over time for the IKDC form, the Lysholm score, and the Cincinnati knee score. A prospective study of 120 ACL reconstructed patients with a 2-year follow-up. Knee Surg Sports Traumatol Arthrosc. 1999;7:152-9.

5. Sernert N, Kartus J, Köhler K, Stener S, Larsson J, Eriksson BI, Karlsson J. Analysis of subjective, objective and functional examination tests after anterior cruciate ligament reconstruction. A follow-up of 527 patients. Knee Surg Sports Traumatol Arthrosc. 1999;7:160-5.

6. Wright RW, Baumgarten KM. Shoulder outcomes measures. J Am Acad Orthop Surg. 2010;18:436-44.

7. Wright RW. Knee injury outcomes measures. J Am Acad Orthop Surg. 2009;17:31-9.

8. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10:407-15.

9. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr Opin Rheumatol. 2002;14:109-14.

10. Martin DP, Engelberg R, Agel J, Snapp D, Swiontkowski MF. Development of a musculoskeletal extremity health status instrument: the Musculoskeletal Function Assessment instrument. J Orthop Res. 1996;14:173-81.

11. Martin DP, Engelberg R, Agel J, Swiontkowski MF. Comparison of the Musculoskeletal Function Assessment questionnaire with the Short Form-36, the Western Ontario and McMaster Universities Osteoarthritis Index, and the Sickness Impact Profile health-status measures. J Bone Joint Surg Am. 1997;79:1323-35.

12. Swiontkowski MF, Engelberg R, Martin DP, Agel J. Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness. J Bone Joint Surg Am. 1999;81:1245-60.

13. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473-83.

14. 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;34:220-33.

15. 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;15:1833-40.

16. Likert R. A technique for the measurement of attitudes. Arch Psychol. 1932;140:1-55.

17. Guyatt GH, Townsend M, Berman LB, Keller JL. A comparison of Likert and visual analogue scales for measuring change in function. J Chronic Dis. 1987;40:1129-33.

18. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum. 2001;45:384-91.

19. Ryser L, Wright BD, Aeschlimann A, Mariacher-Gehler S, Stucki G. A new look at the Western Ontario and McMaster Universities Osteoarthritis Index using Rasch analysis. Arthritis Care Res. 1999;12:331-5.

20. Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis. 1999;58:563-8.

21. Bae SC, Lee HS, Yun HR, Kim TH, Yoo DH, Kim SY. Cross-cultural adaptation and validation of Korean Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research. Osteoarthritis Cartilage. 2001;9:746-50.

22. Bellamy N. WOMAC: a 20-year experiential review of a patient-centered self-reported health status questionnaire. J Rheumatol. 2002;29:2473-6.

23. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. Clin Rheumatol. 2002;21:466-71.

24. Guermazi M, Poiraudeau S, Yahia M, Mezganni M, Fermanian J, Habib Elleuch M, Revel M. Translation, adaptation and validation of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for an Arab population: the Sfax modified WOMAC. Osteoarthritis Cartilage. 2004;12:459-68.

25. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). J Orthop Sci. 2003;8:288-93.

26. Roos EM, Klässbo M, Lohmander LS. WOMAC Osteoarthritis Index. Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and MacMaster Universities. Scand J Rheumatol. 1999; 28:210-5.

27. Salaffi F, Leardini G, Canesi B, Mannoni A, Fioravanti A, Caporali R, Lapadula G, Punzi L; GOnorthrosis and Quality Of Life Assessment (GOQOLA). Reliability and validity of the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in Italian patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2003;11:551-60.

28. Wigler I, Neumann L, Yaron M. Validation study of a Hebrew version of WOMAC in patients with osteoarthritis of the knee. Clin Rheumatol. 1999;18:402-5.

29. Söderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? Clin Orthop Relat Res. 2001;384:189-97.

30. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol. 2000;27:2635-41.

31. Suk M, Hanson BP, Norvell DC, Helfet DL. Musculoskeletal outcomes measures and instruments. AO Handbook. Vol 1. New York: Thieme; 2005.

32. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. J Bone Joint Surg Am. 1969;51:737-55.

33. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol. 1997;50:239-46.

34. Hoeksma HL, Van Den Ende CH, Ronday HK, Heering A, Breedveld FC. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip. Ann Rheum Dis. 2003;62:935-8.

35. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. Arthroscopy. 2000;16:578-87.

36. Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)—validity and responsiveness in total hip replacement. BMC Musculoskelet Disord. 2003;4:10.

37. Martin RL, Philippon MJ. Evidence of validity for the Hip Outcome Score in hip arthroscopy. Arthroscopy. 2007;23:822-6.

38. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the Hip Outcome Score. Arthroscopy. 2006;22:1304-11.

39. Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the Hip Outcome Score. Arthroscopy. 2008;24:676-82.

40. Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. Clin Orthop Relat Res. 2003;406:75-83.

41. Amstutz HC, Thomas BJ, 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;66:228-41.

42. Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. J Arthroplasty. 1998;13:890-5.

43. Roos EM, Roos HP, Lohmander LS. WOMAC Osteoarthritis Index—additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and MacMaster Universities. Osteoarthritis Cartilage. 1999;7:216-21.

44. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003;1:64.

45. 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.

46. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. Arthritis Rheum. 2003;48:2178-87.

47. Roos EM, Ostenberg A, Roos H, Ekdahl C, Lohmander LS. Long-term outcome of meniscectomy: symptoms, function, and performance tests in patients with or without radiographic osteoarthritis compared to matched controls. Osteoarthritis Cartilage. 2001;9:316-24.

48. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)—validation of a Swedish version. Scand J Med Sci Sports. 1998;8:439-48.

49. Roos EM, Roos HP, Ryd L, Lohmander LS. Substantial disability 3 months after arthroscopic partial meniscectomy: a prospective study of patient-relevant outcomes. Arthroscopy. 2000;16:619-26.

50. W-Dahl A, Toksvig-Larsen S, Roos EM. A 2-year prospective study of patient-relevant outcomes in patients operated on for knee osteoarthritis with tibial osteotomy. BMC Musculoskelet Disord. 2005;6:18.

51. Irrgang 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;29:600-13.

52. Irrgang JJ, Anderson AF. Development and validation of health-related quality of life measures for the knee. Clin Orthop Relat Res. 2002;402:95-109.

53. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Neyret P, Richmond JC, Shelbourne KD; International Knee Documentation Committee. Responsiveness of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med. 2006;34:1567-73.

54. Greco NJ, Anderson AF, Mann BJ, Cole BJ, Farr J, Nissen CW, Irrgang JJ. Responsiveness of the International Knee Documentation Committee Subjective Knee Form in comparison to the Western Ontario and McMaster Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal articular cartilage defects. Am J Sports Med. 2010;38:891-902.

55. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res. 1989;248:13-4.

56. König A, Scheidler M, Rader C, Eulert J. The need for a dual rating system in total knee arthroplasty. Clin Orthop Relat Res. 1997;345:161-7.

57. Kreibich DN, Vaz M, Bourne RB, Rorabeck CH, Kim P, Hardie R, Kramer J, Kirkley A. What is the best way of assessing outcome after total knee replacement? Clin Orthop Relat Res. 1996;331:221-5.

58. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med. 1982;10:150-4.

59. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res. 1985;198:43-9.

60. Lukianov AV, Gillquist J, Grana WA, DeHaven KE. An anterior cruciate ligament (ACL) evaluation format for assessment of artificial or autologous anterior cruciate reconstruction results. Clin Orthop Relat Res. 1987;218:167-80.

61. Bengtsson J, Möllborg J, Werner S. A study for testing the sensitivity and reliability of the Lysholm knee scoring scale. Knee Surg Sports Traumatol Arthrosc. 1996;4:27-31.

62. Irrgang JJ, Ho H, Harner CD, Fu FH. Use of the International Knee Documentation Committee guidelines to assess outcome following anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc. 1998;6:107-14.

63. Noyes FR, Matthews DS, Mooar PA, Grood ES. The symptomatic anterior cruciate-deficient knee. Part II: the results of rehabilitation, activity modification, and counseling on functional disability. J Bone Joint Surg Am. 1983;65:163-74.

64. Noyes FR, McGinniss GH, Mooar LA. Functional disability in the anterior cruciate insufficient knee syndrome. Review of knee rating systems and projected risk factors in determining treatment. Sports Med. 1984;1:278-302.

65. Barber-Westin SD, Noyes FR, McCloskey JW. Rigorous statistical reliability, validity, and responsiveness testing of the Cincinnati Knee Rating System in 350 subjects with uninjured, injured, or anterior cruciate ligament-reconstructed knees. Am J Sports Med. 1999;27:402-16.

66. Bollen S, Seedhom BB. A comparison of the Lysholm and Cincinnati knee scoring questionnaires. Am J Sports Med. 1991;19:189-90.

67. Sgaglione NA, Del Pizzo W, Fox JM, Friedman MJ. Critical analysis of knee ligament rating systems. Am J Sports Med. 1995;23:660-7.

68. Marx RG, Jones EC, Allen AA, Altchek DW, O'Brien SJ, Rodeo SA, Williams RJ, Warren RF, Wickiewicz TL. Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. J Bone Joint Surg Am. 2001;83: 1459-69.

69. Mohtadi N. Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency. Am J Sports Med. 1998;26:350-9.

70. Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and evaluation of an activity rating scale for disorders of the knee. Am J Sports Med. 2001;29:213-8.

71. Suk M, Hanson BP, Norvell DC, Helfet DL. Musculoskeletal outcomes measures and instruments. AO Handbook. Vol 1 and 2. New York: Thieme; 2009.

72. André M, Hagelberg S, Stenström CH. The juvenile arthritis foot disability index: development and evaluation of measurement properties. J Rheumatol. 2004; 31:2488-93.

73. Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. J Clin Epidemiol. 1991;44:561-70.

74. Helliwell P, Reay N, Gilworth G, Redmond A, Slade A, Tennant A, Woodburn J. Development of a foot impact scale for rheumatoid arthritis. Arthritis Rheum. 2005; 53:418-22.

75. Kuyvenhoven MM, Gorter KJ, Zuithoff P, Budiman-Mak E, Conrad KJ, Post MW. The Foot Function Index with verbal rating scales (FFI-5pt): a clinimetric evaluation and comparison with the original FFI. J Rheumatol. 2002;29:1023-8.

76. Saag KG, Saltzman CL, Brown CK, Budiman-Mak E. The Foot Function Index for measuring rheumatoid arthritis pain: evaluating side-to-side reliability. Foot Ankle Int. 1996;17:506-10.

77. Halasi T, Kynsburg A, Tállay A, Berkes I. Development of a new activity score for the evaluation of ankle instability. Am J Sports Med. 2004;32:899-908.

78. Hale SA, Hertel J. Reliability and sensitivity of the Foot and Ankle Disability Index in subjects with chronic ankle instability. J Athl Train. 2005;40:35-40.

79. Liu SH, Jacobson KE. A new operation for chronic lateral ankle instability. J Bone Joint Surg Br. 1995;77:55-9.

80. Pugia ML, Middel CJ, Seward SW, Pollock JL, Hall RC, Lowe L, Mahony L, Henderson NE. Comparison of acute swelling and function in subjects with lateral ankle injury. J Orthop Sports Phys Ther. 2001;31:384-8.

81. Roos EM, Brandsson S, Karlsson J. Validation of the foot and ankle outcome score for ankle ligament reconstruction. Foot Ankle Int. 2001;22:788-94.

82. Rozzi SL, Lephart SM, Sterner R, Kuligowski L. Balance training for persons with functionally unstable ankles. J Orthop Sports Phys Ther. 1999;29:478-86.

83. Williams GN, Molloy JM, DeBerardino TM, Arciero RA, Taylor DC. Evaluation of the Sports Ankle Rating System in young, athletic individuals with acute lateral ankle sprains. Foot Ankle Int. 2003;24:274-82.

84. Nilsson-Helander K, Thomeé R, Silbernagel KG, Thomeé P, Faxén E, Eriksson BI, Karlsson J. The Achilles tendon Total Rupture Score (ATRS): development and validation. Am J Sports Med. 2007;35:421-6.

85. Robinson JM, Cook JL, Purdam C, Visentini PJ, Ross J, Maffulli N, Taunton JE, Khan KM; Victorian Institute Of Sport Tendon Study Group. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. Br J Sports Med. 2001;35:335-41.

86. Fuchs S, Sandmann C, Skwara A, Chylarecki C. Quality of life 20 years after arthrodesis of the ankle. A study of adjacent joints. J Bone Joint Surg Br. 2003; 85:994-8.

87. Kofoed H, Sørensen TS. Ankle arthroplasty for rheumatoid arthritis and osteoarthritis: prospective long-term study of cemented replacements. J Bone Joint Surg Br. 1998;80:328-32.

88. Domsic RT, Saltzman CL. Ankle Osteoarthritis Scale. Foot Ankle Int. 1998; 19:466-71.

89. Sanders R, Fortin P, DiPasquale T, Walling A. Operative treatment in 120 displaced intraarticular calcaneal fractures. Results using a prognostic computed tomography scan classification. Clin Orthop Relat Res. 1993;290:87-95.

90. Thordarson DB, Krieger LE. Operative vs. nonoperative treatment of intraarticular fractures of the calcaneus: a prospective randomized trial. Foot Ankle Int. 1996;17:2-9.

91. Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. J Bone Joint Surg Am. 2004;86:902-9.

92. Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M. Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. Foot Ankle Int. 1994;15:349-53.

93. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). Foot Ankle Int. 2005;26:968-83.

94. Ibrahim T, Beiri A, Azzabi M, Best AJ, Taylor GJ, Menon DK. Reliability and validity of the subjective component of the American Orthopaedic Foot and Ankle Society clinical rating scales. J Foot Ankle Surg. 2007;46:65-74.

95. Malviya A, Makwana N, Laing P. Correlation of the AOFAS scores with a generic health QUALY score in foot and ankle surgery. Foot Ankle Int. 2007;28: 494-8.

96. SooHoo NF, Shuler M, Fleming LL; American Orthopaedic Foot and Ankle Society. Evaluation of the validity of the AOFAS Clinical Rating Systems by correlation to the SF-36. Foot Ankle Int. 2003;24:50-5.

97. Westphal T, Piatek S, Halm JP, Schubert S, Winckler S. Outcome of surgically treated intraarticular calcaneus fractures—SF-36 compared with AOFAS and MFS. Acta Orthop Scand. 2004;75:750-5.

98. Button G, Pinney S. A meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system? Foot Ankle Int. 2004;25: 521-5.

99. Guyton GP. Theoretical limitations of the AOFAS scoring systems: an analysis using Monte Carlo modeling. Foot Ankle Int. 2001;22:779-87.

100. Baumhauer JF, Nawoczenski DA, DiGiovanni BF, Wilding GE. Reliability and validity of the American Orthopaedic Foot and Ankle Society Clinical Rating Scale: a pilot study for the hallux and lesser toes. Foot Ankle Int. 2006;27:1014-9.

101. SooHoo NF, Vyas R, Samimi D. Responsiveness of the foot function index, AOFAS clinical rating systems, and SF-36 after foot and ankle surgery. Foot Ankle Int. 2006;27:930-4.

102. Carcia CR, Martin RL, Drouin JM. Validity of the Foot and Ankle Ability Measure in athletes with chronic ankle instability. J Athl Train. 2008;43:179-83.

103. Martin RL, Hutt DM, Wukich DK. Validity of the Foot and Ankle Ability Measure (FAAM) in diabetes mellitus. Foot Ankle Int. 2009;30:297-302.

104. Agel J, Beskin JL, Brage M, Guyton GP, Kadel NJ, Saltzman CL, Sands AK, Sangeorzan BJ, SooHoo NF, Stroud CC, Thordarson DB. Reliability of the Foot Function Index: a report of the AOFAS Outcomes Committee. Foot Ankle Int. 2005; 26:962-7.

105. SooHoo NF, Samimi DB, Vyas RM, Botzler T. Evaluation of the validity of the Foot Function Index in measuring outcomes in patients with foot and ankle disorders. Foot Ankle Int. 2006;27:38-42.

106. Budiman-Mak E, Conrad K, Stuck R, Matters M. Theoretical model and Rasch analysis to develop a revised Foot Function Index. Foot Ankle Int. 2006;27:519-27.

107. Knecht SI, Estin M, Callaghan JJ, Zimmerman MB, Alliman KJ, Alvine FG, Saltzman CL. The Agility total ankle arthroplasty. Seven to sixteen-year follow-up. J Bone Joint Surg Am. 2004;86:1161-71.

108. Thomas R, Daniels TR, Parker K. Gait analysis and functional outcomes following ankle arthrodesis for isolated ankle arthritis. J Bone Joint Surg Am. 2006;88: 526-35.

109. Marsh JL, McKinley T, Dirschl D, Pick A, Haft G, Anderson DD, Brown T. The sequential recovery of health status after tibial plafond fractures. J Orthop Trauma. 2010;24:499-504.

110. Marsh JL, Muehling V, Dirschl D, Hurwitz S, Brown TD, Nepola J. Tibial plafond fractures treated by articulated external fixation: a randomized trial of postoperative motion versus nonmotion. J Orthop Trauma. 2006;20:536-41.

111. Mazur JM, Schwartz E, Simon SR. Ankle arthrodesis. Long-term follow-up with gait analysis. J Bone Joint Surg Am. 1979;61:964-75.

112. Abdo RV, Wasilewski SA. Ankle arthrodesis: a long-term study. Foot Ankle. 1992;13:307-12.

113. Boraiah S, Kemp TJ, Erwteman A, Lucas PA, Asprinio DE. Outcome following open reduction and internal fixation of open pilon fractures. J Bone Joint Surg Am. 2010;92:346-52.

114. Kopp FJ, Banks MA, Marcus RE. Clinical outcome of tibiotalar arthrodesis utilizing the chevron technique. Foot Ankle Int. 2004;25:225-30.

115. Percy EC, Conochie LB. The surgical treatment of ruptured tendo Achillis. Am J Sports Med. 1978;6:132-6.

116. Curwin S, Stanish WD. Tendinitis: its etiology and treatment. Lexington: Collamore Press; 1984.

117. Silbernagel KG, Thomeé R, Karlsson J. Cross-cultural adaptation of the VISA-A questionnaire, an index of clinical severity for patients with Achilles tendinopathy, with reliability, validity and structure evaluations. BMC Musculoskelet Disord. 2005;6:12.

118. Brauer CA, Neumann PJ, Rosen AB. Trends in cost effectiveness analyses in orthopaedic surgery. Clin Orthop Relat Res. 2007;457:42-8.