HHS Public Access

Author manuscript

Rheum Dis Clin North Am. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Rheum Dis Clin North Am. 2016 May ; 42(2): 239–252. doi:10.1016/j.rdc.2016.01.005.

Patient Reported Measures of Physical Function in Knee Osteoarthritis

Daniel K. White, PT, ScD, MSc[1] and **Hiral Master, DPT, MPH**[1]

[1]Department of Physical Therapy, University of Delaware, Newark, DE

Summary

Knee osteoarthritis (OA) is a common cause of an array of functional limitations in older adults, and the accurate assessment of such limitations is critical for the practicing clinician and scientist. Patient reported measures are a valuable resource to track the type and severity of limitation, though the psychometric performance of each instrument should be thoroughly evaluated before adoption. We review the validity, reliability, sensitivity to change, and responsiveness of three patient reported measures of physical function: The Western Ontario and McMasters University Osteoarthritis Index (WOMAC), the Knee Injury and Osteoarthritis Outcome Score (KOOS), and the Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function scale.

Introduction

Knee osteoarthritis (OA) is a leading cause of functional limitation worldwide.[1,2] People with knee OA have pain, which limits commonly performed daily activities. Functional limitation is defined by Nagi as restriction in the performance of an individual, such as difficulty getting up out of bed, getting up from a chair, walking, and climbing stairs.[3] Functional limitation is a construct that is unique and separate from impairments, e.g., knee pain, and disease, e.g., knee OA.

Accurately assessing the type and severity of functional limitation is important for people with knee OA. From a societal perspective, proper measurement helps determine the burden of disease on function. From a research perspective, evaluating the efficacy and effectiveness of new treatment interventions requires measurement of physical function using appropriate measures. Lastly, from a clinical perspective, assessing functional limitation is important to demonstrate the efficacy of one-on-one intervention and describe worsening or improvement over time.[4] Measurement in knee OA is challenging since there is a wide spectrum of functional limitation types and severity.

Corresponding author: Daniel K. White PT, ScD, MSc, 540 South College Ave, 210L, Newark, DE 19713, P: (302) 831-7607, F: (302) 831-4234, dkw@udel.edu.

The authors have no conflicts of interest

Examples of commonly employed fixed length questionnaires of physical function in knee OA include the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), the Knee injury and Osteoarthritis Outcome Score (KOOS), and the Patient Reported Outcomes Measurement Information System (PROMIS) physical function subscale. WOMAC is a disease specific instrument that measures the domains of pain, stiffness, and physical function.[5] For purposes of this paper, we will focus on the physical function subscale. The KOOS another disease specific instrument that has similar items as the WOMAC physical function subscale with the addition of questions about sport and recreation and knee related Quality of life (QOL).[6,7] The Patient-Reported Outcomes Measurement Information System (PROMIS) physical function instrument is recently developed general measure of health.[8]

The purpose of this paper is to review psychometric properties of commonly used patient reported measures of physical function in knee OA, and discuss the strengths and limitations of each measurement instrument.

The importance of measuring the construct of physical function

The disablement model is a useful tool to communicate the consequences of injury and disease. Jette adopted Nagi's definition of disablement as "various impacts of chronic and acute conditions on the functioning of specific body systems, on basic human performance, and on people's functioning in necessary, usual, expected, and personally desired roles in society."[9] Physical function has a unique place in contemporary disablement frameworks. Functional limitation is a distinct phenomena in the Nagi model that describes the construct of physical function, defined as restriction in the performance of an individual, such as difficulty getting up out of bed, getting up from a chair, walking, and climbing stairs.[3] In the more recent International Classification of Function, Disability, and Health (ICF) model from the World Health Organization, Activity and Activity Limitation best describe physical function. Activity is "… the execution of a task or action by an individual. Activity limitations are defined as difficulties an individual may have in executing activities."[10]

From a measurement perspective, outcome instruments should measure specific underlying constructs and not mix two or more constructs together. For instance, while physical function is closely related to disease and pain, it is a unique construct of disablement. To best understand the prevalence and associated risk factors of functional limitation, an ideal measurement instrument should attempt to solely measure the construct of physical function.

Psychometric Properties

We reviewed the following psychometric properties of patient reported measures: reliability, validity, and sensitivity to change and responsiveness (Table 1). We also investigated known group validity, which is the extent to which a measurement instrument can differentiate scores from groups that are known to be different. Lastly, we evaluated sensitivity to change and responsiveness. Sensitivity to change is the ability of measure to detect change that exceeds statistical error without regard to clinical relevance, while responsiveness refers to clinically relevant or meaningful change.[13] Responsiveness is determined using scores

anchored to patient or provider reported thresholds, such as the Minimum Clinical Important Difference (MCID).[13,14]

The greater number of psychometric properties studied and properly fulfilled, the better a measurement instrument will assess physical function. It is important for clinicians and researchers to be familiar with the psychometric properties of commonly employed measurement instruments in order to help guide clinical decision-making and appropriately balance the implications of study findings.

Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC) Physical function subscale

The WOMAC physical function subscale is a widely employed patient reported measurement instrument for knee OA. The WOMAC was developed in order to fill the need for an outcome instrument that could be responsive to change in OA-related symptoms following a clinical trial.[5] Bellamy and Buchanan interviewed 100 people with knee OA and identified 41 items in five dimensions that were important for people with knee OA, one being physical function.[15] The WOMAC has been used in other patient populations including hip OA,[16] rheumatoid arthritis and fibromyalgia,[17] and Systemic Lupus Erythematosus[16] among others.

The WOMAC physical function subscale beings with the question, “What degree of difficulty do you have due to pain, discomfort or arthritis…”. Seventeen items are evaluated including going up and down stairs, sitting, standing, squatting to the floor, walking, getting in a car, shopping, putting on and taking off socks, getting out of bed, lying in bed, bathing, sitting, getting on and off the toilet, heavy domestic duties, and light domestic duties. There are two methods of scoring. The Likert method uses a five-point scale with the choices of none, mild, moderate, severe, or extreme and the physical function subscale ranges from 0 to 68. The VAS method uses a 100mm horizontal line for each item and subjects mark a vertical line along the horizontal continuum for each item, which is measured and totaled with the other items on a 0 to 1700 scale.[18] For both scoring methods, higher scores represent more functional limitation.

Reliability

The WOMAC physical function subscale shows good reliability. Specifically, internal consistency of the WOMAC physical function subscale is high in English and other languages. As well, high test-reliability over durations of time ranging from 6 days to 12 months has been shown as well. (Table 2)

Validity

The WOMAC physical function subscale shows good construct validity with other known measures of physical function. (Table 2) As well, the WOMAC physical function subscale has been shown to have negligible flooring and ceiling effects.[19] We did not find studies that validated the WOMAC to known-groups.

Sensitivity to change and Responsiveness

Previous studies have shown the WOMAC physical function subscale to be as sensitive to change as other measures of physical function. Responsiveness thresholds have been established as well. (Table 4)

Strengths and Weaknesses

The major strengths of the WOMAC physical function subscale are that this measurement instrument has been well validated, has good test-retest reliability, and has established MCID thresholds. Furthermore, since the instrument was developed as a disease specific outcome, people with knee OA are likely to have difficulty with the items from the scale. Lastly, the instrument is easy to administer and has been translated into many different languages.

One potential major limitation with the WOMAC physical function subscale is its unclear delineation between the constructs of pain and function. This is because each item begins with the question “What degree of difficulty do you have due to pain, discomfort or arthritis?” Hence, respondents technically only report the extent of functional limitation attributed to pain, discomfort or arthritis, which may not reflect their total limitation. One previous study found the WOMAC physical function to a stronger association with measures of pain compared with performance measures of physical function.[20] Investigators should consider employing an additional measure of physical function to accompany the WOMAC to fully measure the construct of physical function.

Knee Injury and Osteoarthritis Outcome Score (KOOS): Function, Sports, and Recreational Activity, and KOOS Physical Function Short Form (KOOS-PS)

Ewa Roos and colleagues developed the KOOS using the WOMAC, a literature review, an expert panel of patients, orthopedic surgeons, and physical therapists, and data from a pilot study of people with post-traumatic OA.[7] An important innovation of the KOOS over the WOMAC is the addition of a high functioning subscale termed ‘Function, Sports, and Recreational Activity’. Respondents assess the degree of difficulty experienced during the last week due to their knee in five items, Squatting, Running, Jumping, Twisting/pivoting on the injured knee and kneeling. Responses are on a Likert scale ranging from ‘None’ to ‘Extreme’.

KOOS-Physical Function (KOOS-PS) Short Form is a parsimonious measure of physical function derived from the KOOS. A working group tasked with constructing a composite measure of OA severity sponsored by the Osteoarthritis Research Society International (OARSI) and Outcome Measures in Rheumatology (OMERACT) developed the KOOS-PS.[21] Respondents are asked to indicate the degree of difficulty experienced in the last week due to a knee problem in 7 tasks: Rising from Bed, Putting on socks/stockings, Rising from sitting, Bending to the floor, Twisting/pivoting on the injured knee, Kneeling, and Squatting. Responses are on a Likert scale ranging from ‘None’ to ‘Extreme’.

Reliability

The KOOS Function, Sports, and Recreational Activity subscale shows good internal consistency and test-retest reliability in multiple studies. The measurement instrument has good reported test-retest reliability measured over durations of time ranging from 2 to 14 days. (Table 4) The KOOS-Physical Function Short Form also has good internal consistency and high test-retest reliability, which was measured over durations of time ranging from 2 to 14 days.

Validity

The Function, Sports, and Recreational Activity subscale were correlated with other known measures of physical function, namely the SF-36 Physical function subscale. We did find several studies reporting floor effects.[22–24] Statistically significant differences in scores were found for known groups, e.g., severity of knee OA[23], treatment approach,[25] and use of walking aids.[24] Several studies reported the KOOS Function, Sports, and Recreational Activity subscale to be sensitive to change,[24,25] though we found that less is known regarding responsiveness. For the KOOS-Physical Function Short Form, we found high correlations with established measures of physical function and no reported floor or ceiling effects. One study reported differences in scores for people with symptomatic knee OA with and without assistive device use.[26] (Table 6)

Sensitivity to change and Responsiveness

The Function, Sports, and Recreational Activity subscale was sensitive to change, although little is known about responsiveness. The KOOS-PS short form was also sensitive to change, and one study established an MCID. (Table 5)

Strengths and Weaknesses

The KOOS Function, Sports, and Recreational Activity subscale and KOOS-PS Short Form have several strengths. First, both instruments are well studied from across the globe and are reliable, valid, and sensitive to change. Second, the KOOS is freely available by download http://www.koos.nu in different languages. Third, the KOOS is disease specific to knee OA and builds on the strengths of the WOMAC. Lastly, the KOOS-PS Short Form is brief and takes minimal time to administer.

There are several limitations to the KOOS. First, the Function, Sports, and Recreational Activity subscale has reported floor effects with people with low functional status. This may not be unexpected given this subscale was developed to measure higher level functional ability. Second, little is known about responsiveness of the KOOS. While the KOOS is sensitive to change, few studies reported thresholds needed to reach meaningful change, whether by a distribution based (MDC) or anchor based methodology (MCID).

Patient Reported Outcomes Measurement Information System (PROMIS) Physical Function

PROMIS was created in response to the need to speed research discoveries by providing a national resource for accurate and efficient measurement of patient-reported outcomes. The initiative, funded under the National Institutes of Health (NIH) Roadmap for Medical Research, sought to improve the reliability, validity, and precision of patient-reported outcomes and create new measurements that exceeded the psychometric performance of legacy measures.[27] Physical function was included as a domain included within the PROMIS framework.

We summarize the steps used to develop the PROMIS physical function domain, which is described online at http://www.nihpromis.org/Documents/PROMIS_The_First_Four_Years.pdf. As a first step, an item bank was developed using both qualitative and quantitative methods. Items from established instruments were included in the item bank. Next, items were subjected to Item-response theory (IRT) analyses in order to better understand their dimensional structure. The number of items that were the most representative and informative were chosen through placing common items in a similar 'bin' and next 'winnowing' out items that were redundant or of less quality to alternative items in the same bin. The wording of remaining items was reviewed and revised by an expert panel. This was followed by focus groups to inform definitions of the domains and identify areas of future development and cognitive interviews to examine individual item comprehension.[28] Reviewed item banks were then field tested in the general population and specific patient populations, including people with arthritis. The first wave of testing included 7000 people demographic that was similar to the US census in 2000.

Presently, PROMIS instruments are freely available online at http://www.assessmentcenter.net/. There are five physical function domain PROMIS instruments relevant to people with knee OA with the number of items ranging from 4 to 20. Response choices are in a Likert format. For items asking "Does your health now limit you in …" response choices range from "Not at all" to "Cannot do". For items asking "Are you able to …" response choice range from "Without any difficulty" to "Unable to do". For each item, the item statistics are provided with the mean score and frequency of response choices using data from the first wave of testing.

Reliability

There were few studies to date that examined reliability among people with knee OA. We did find high internal consistency for people with RA and Spanish-speaking adults. We also found one study that reported high test-retest reliability over 4 weeks among people with doctor- diagnosed OA. (Table 6)

Validity

We found one study that showed high correlation of PROMIS physical function scores were found with SF-36 scores. PROMIS scores were also highly correlated with physical function measures in people with RA. (Table 7) Among people with RA, the PROMIS Physical

Function instrument covers a wider range of physical function levels than the Health Assessment Questionnaire- Disability Index.[29] PROMIS also has worse scores with increased disease activity,[30] and lower scores between those in remission compared with those with active RA disease.[29] Among 204 people with knee OA, almost none had floor or ceiling effects.[31]

Sensitivity to change and Responsiveness

Little is known about sensitivity to change and responsiveness for PROMIS in knee OA. One study in RA reported the PROMIS Physical Function 20-item instrument to be more sensitive to change than the PF-10.[32]

Strengths and Weaknesses

A major strength of PROMIS physical function is the rigorous development of the item bank that yielded a conceptually clear and well-calibrated measurement instrument. This laid the groundwork for development of a Computer Adapted Test (CAT) format of the PROMIS physical function.. Another strength is that all the PROMIS measures are freely available at http://www.assessmentcenter.net/.

Given the relatively recent development of PROMIS physical function, literature demonstrating its reliability, validity, sensitivity to change, and responsiveness is a work in progress. To date, psychometric studies of PROMIS measures have shown positive results, although more work is needed.

Summary

Adequately assessing the burden of knee OA on physical function is challenging given the large number of functional limitations possible in this patient population. The WOMAC and KOOS are useful instruments that provide a valid and reliable measurement of functional limitation common in knee OA. PROMIS physical function is a newer measurement instrument that has a growing number of publications demonstrating its legitimacy in knee OA.

Acknowledgments

Funding for the authors was provided by NIH U54 GM104941

References

1. Cross M, Smith E, Hoy D, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. Annals of the rheumatic diseases. Jul; 2014 73(7):1323–1330. [PubMed: 24553908]
2. Guccione AA, Felson DT, Anderson JJ, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. American journal of public health. Mar; 1994 84(3):351–358. [PubMed: 8129049]
3. Jette AM. Toward a common language for function, disability, and health. Physical therapy. May; 2006 86(5):726–734. [PubMed: 16649895]
4. White D, Neogi T, Nguyen US, Niu J, Zhang Y. Trajectories of functional decline in knee osteoarthritis: The OAI. Rheumatology (Oxford). 2015 In Press.

5. 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. The Journal of rheumatology. Dec; 1988 15(12):1833–1840. [PubMed: 3068365]
6. 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. Dec; 1998 8(6):439–448. [PubMed: 9863983]
7. 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. Aug; 1998 28(2):88–96. [PubMed: 9699158]
8. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). Journal of clinical epidemiology. Jan; 2008 61(1):17–33. [PubMed: 18083459]
9. Jette AM. Physical disablement concepts for physical therapy research and practice. Physical therapy. May; 1994 74(5):380–386. [PubMed: 8171099]
10. International Classification of Functioning, Disability, and Health: ICF. Geneva, Switzerland: World Health Organization; 2001.
11. Furr, RM.; Bacharach, VR. Psychometrics : an introduction. 2. Lod Angeles: SAGE; 2014.
12. Hays, RD.; Revicki, D. Reliability and validity (including responsiveness). Oxford: Oxford University Press; 1988.
13. Latham NK, Mehta V, Nguyen AM, et al. Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? Arch Phys Med Rehabil. Nov; 2008 89(11):2146–2155. [PubMed: 18996244]
14. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. The Journal of rheumatology. Feb; 2001 28(2):400–405. [PubMed: 11246687]
15. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol. Jun; 1986 5(2):231–241. [PubMed: 3731718]
16. Stucki G, Sangha O, Stucki S, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the selfadministered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage. Mar; 1998 6(2):79–86. [PubMed: 9692062]
17. 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. Sep; 1999 58(9):563–568. [PubMed: 10460190]
18. Bruce B, Fries J. Longitudinal comparison of the Health Assessment Questionnaire (HAQ) and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Arthritis Rheum. Oct 15; 2004 51(5):730–737. [PubMed: 15478152]
19. Tuzun EH, Eker L, Aytar A, Daskapan A, Bayramoglu M. Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthritis Cartilage. Jan; 2005 13(1):28–33. [PubMed: 15639634]
20. Terwee CB, van der Slikke RM, van Lummel RC, Benink RJ, Meijers WG, de Vet HC. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. Journal of clinical epidemiology. Jul; 2006 59(7):724–731. [PubMed: 16765276]
21. Perruccio AV, Stefan Lohmander L, Canizares M, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. Osteoarthritis Cartilage. May; 2008 16(5):542–550. [PubMed: 18294869]
22. Xie F, Li SC, Roos EM, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with

knee osteoarthritis in Singapore. Osteoarthritis Cartilage. Nov; 2006 14(11):1098–1103. [PubMed: 16814575]

23. de Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. Health and quality of life outcomes. 2008; 6:16. [PubMed: 18302729]
24. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score (KOOS). Osteoarthritis Cartilage. Sep; 2009 17(9):1156–1162. [PubMed: 19303082]
25. Ornetti P, Parratte S, Gossec L, et al. Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. Osteoarthritis Cartilage. Apr; 2008 16(4):423–428. [PubMed: 17905602]
26. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL, Gil J. Reliability, validity and responsiveness of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score--Physical Function Short-form (KOOS-PS). Osteoarthritis Cartilage. Mar; 2010 18(3):372–376. [PubMed: 19912982]
27. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol. Sep-Oct;2005 23(5 Suppl 39):S53–57. [PubMed: 16273785]
28. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. Med Care. May; 2007 45(5 Suppl 1):S12–21. [PubMed: 17443114]
29. Oude Voshaar MA, Ten Klooster PM, Glas CA, et al. Validity and measurement precision of the PROMIS physical function item bank and a content validity-driven 20-item short form in rheumatoid arthritis compared with traditional measures. Rheumatology (Oxford). Jul 29.2015
30. Bartlett SJ, Orbai AM, Duncan T, et al. Reliability and Validity of Selected PROMIS Measures in People with Rheumatoid Arthritis. PLoS One. 2015; 10(9):e0138543. [PubMed: 26379233]
31. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. BMC Musculoskelet Disord. 2015; 16:253. [PubMed: 26369412]
32. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. Arthritis Res Ther. 2011; 13(5):R147. [PubMed: 21914216]
33. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. Qual Life Res. 2007; 16(Suppl 1):133–141. [PubMed: 17401637]
34. Jette AM, Haley SM, Ni P, Olarsch S, Moed R. Creating a computer adaptive test version of the late-life function and disability instrument. The journals of gerontology. Nov; 2008 63(11):1246–1256. [PubMed: 19038841]
35. Kosinski M, Bjorner JB, Ware JE Jr, Sullivan E, Straus WL. An evaluation of a patient-reported outcomes found computerized adaptive testing was efficient in assessing osteoarthritis impact. J Clin Epidemiol. Jul; 2006 59(7):715–723. [PubMed: 16765275]
36. McDonough CM, Tian F, Ni P, et al. Development of the computer-adaptive version of the Late-Life Function and Disability Instrument. J Gerontol A Biol Sci Med Sci. Dec; 2012 67(12):1427–1438. [PubMed: 22546960]
37. Jette AM, McDonough CM, Ni P, et al. A functional difficulty and functional pain instrument for hip and knee osteoarthritis. Arthritis Res Ther. 2009; 11(4):R107. [PubMed: 19589168]
38. Jette AM, McDonough CM, Haley SM, et al. A computer-adaptive disability instrument for lower extremity osteoarthritis research demonstrated promising breadth, precision, and reliability. J Clin Epidemiol. Aug; 2009 62(8):807–815. [PubMed: 19216052]
39. PMFDM. Quality of Life: Assessment, Analysis and Interpretation. John Wiley & Sons; 2002.
40. EFJHGGJR. How to develop and validate a new health-related quality of life instrument. Philadelphia: LIpponcott-Raven Publishers; 1996.
41. Xie F, Li SC, Goeree R, et al. Validation of Chinese Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual Life Res. May; 2008 17(4):595–601. [PubMed: 18415706]

42. Nadrian H, Moghimi N, Nadrian E, et al. Validity and reliability of the Persian versions of WOMAC Osteoarthritis Index and Lequesne Algofunctional Index. Clin Rheumatol. Jul; 2012 31(7):1097–1102. [PubMed: 22526476]
43. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and Mcmaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. Jul; 2001 9(5):440–446.
44. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. Clin Rheumatol. Jul; 2010 29(7):749–756. [PubMed: 20169459]
45. Williams VJ, Piva SR, Irrgang JJ, Crossley C, Fitzgerald GK. Comparison of reliability and responsiveness of patient-reported clinical outcome measures in knee osteoarthritis rehabilitation. J Orthop Sports Phys Ther. Aug; 2012 42(8):716–723. [PubMed: 22402677]
46. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Annals of the rheumatic diseases. Jan; 2005 64(1):29–33. [PubMed: 15208174]
47. Angst F, Aeschlimann A, Steiner W, Stucki G. Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. Ann Rheum Dis. Sep; 2001 60(9):834–840. [PubMed: 11502609]
48. Paradowski PT, Keska R, Witonski D. Validation of the Polish version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in patients with osteoarthritis undergoing total knee replacement. BMJ Open. 2015; 5(7):e006947.
49. Ornetti P, Perruccio AV, Roos EM, Lohmander LS, Davis AM, Maillefert JF. Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS). Osteoarthritis Cartilage. Dec; 2009 17(12):1604–1608. [PubMed: 19589400]
50. Gul ED, Yilmaz O, Bodur H. Reliability and validity of the Turkish version of the knee injury and osteoarthritis outcome score-physical function short-form (KOOS-PS). J Back Musculoskelet Rehabil. 2013; 26(4):461–466. [PubMed: 23948834]
51. Singh JA, Luo R, Landon GC, Suarez-Almazor M. Reliability and clinically important improvement thresholds for osteoarthritis pain and function scales: a multicenter study. J Rheumatol. Mar; 2014 41(3):509–515. [PubMed: 24429183]
52. Davis AM, Perruccio AV, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. Osteoarthritis Cartilage. Jul; 2009 17(7):843–847. [PubMed: 19215728]
53. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and reliability of patient-reported outcomes measurement information system instruments in osteoarthritis. Arthritis Care Res (Hoboken). Oct; 2013 65(10):1625–1633. [PubMed: 23592494]
54. Paz SH, Spritzer KL, Morales LS, Hays RD. Evaluation of the Patient-Reported Outcomes Information System (PROMIS((R))) Spanish-language physical functioning items. Qual Life Res. Sep; 2013 22(7):1819–1830. [PubMed: 23124505]

Key Points

- The WOMAC Physical Function subscale is a well-validated and reliable patient reported measure, although it is questionable if the constructs of physical function and pain are separately evaluated.
- The KOOS: The Function, Sports, and Recreational Activity subscale is a well-validated measure of physical function, though floor effects are present for people with moderate to severe functional limitation.
- The PROMIS Physical Function is a newer measurement instrument, and preliminary studies show high test-retest reliability and no floor or ceiling effects among people with OA.

Table 1

Definitions and implications of statistical values for psychometric testing of patient reported measures

Psychometric Property	Definition	Statistics
Reliability-Internal consistency	Measure the relation of the items in a questionnaire to an underlying construct.	Cronbach's alpha - items are considered to represent a similar construct when alpha is approximately ≥0.7.[12]
Reliability-test–retest or inter/intra-rater	If the questionnaire measures a condition in a reproducible manner.	Intraclass correlation coefficient (ICC). ICC ≥0.7 is considered acceptable for test–retest reliability.[39]
Validity: Construct	Evaluation of the relationship of an instrument with other instruments.	Correlation coefficient (Pearson's r, Spearman's rho). correlation coefficients of >0.50, 0.35–0.50, and <0.35 were considered strong, moderate, and weak, respectively.[40]
Validity: Known groups	Measure of an instrument's ability to distinguish among different groups (e.g., persons with functional limitation versus those without functional limitation).	T-test or analysis of variance (ANOVA), with post-hoc analysis.
Sensitivity to change	Ability to detect change that exceeds statistical error without regard to clinical relevance.	Effect Size (ES), Standardized Response Mean (SRM), and minimal detectable change (MDC). Larger ES and SRM indicates more sensitivity to change.
Responsiveness	Ability to measure clinically relevant or meaningful change.	Minimal Clinically Important Difference

Table 2

Reliability and Validity of the WOMAC physical function subscale

Study	N	Population description	Internal consistency: Cronbach's alpha	Test-retest reliability: ICC	Comparator	Correlation
Bellamy et al 1988[5]	57	People from Canada with knee or hip OA who were taking part in a randomized controlled trial of 2 nonsteroidal anti-inflammatory drugs	Likert Scale format = 0.95. VAS format = 0.89		Lequesne Physical Function	r = 0.5
Xie et al 2008[41]	131	People from Singapore, China with knee OA scheduled for TKA	> 0.7	> 0.8	SF-36 physical functioning and EQ-5D in Chinese and English	r > 0.5
Nadrian et al 2012[42]	116	People from Iran with physician diagnosed referred to a rheumatology clinic with knee or hip OA	0.95	> 0.7	Lequesne Physical Function	r > 0.5
Thumboo et al 2001[43]	66	English-speaking Chinese, Malay or Indian people with knee or hip OA seen at a tertiary referral center	0.93	> 0.8	SF-36 Physical Functioning	r = −0.4
Tüzün, et al., 2005[19]	72	People from Turkey being seen in an outpatient physical therapy practice with knee OA	0.94		SF-36 Physical Functioning, Lequesne Physical Function	r = −0.7, r = 0.7
Basaran, et al., 2010[44]	117	People from Turkey with knee or hip OA	0.95		SF-36 Physical Functioning	r = −0.8
Williams et al 2012[45]	168	People with knee OA participating in a rehabilitation program.		> 0.8		

Table 3

Sensitivity to change and Responsiveness of the WOMAC Physical Function subscale

Study	N	ES and SRM of WOMAC Physical Function vs comparator	MDC 90	MCID
Tubach[46]	1362			Minimally Clinical Important Improvement = 26% change
Angst[47]	223	SRM = 0.63 vs SF-36 Physical Function: SRM = 0.25		
Williams et al 2012[45]	116		Over 12 months: MDC90 change >= 11.8/100	Similar MCID as the Lower extremity functional scale
Tüzün, et al., 2005[19]	72	ES = 0.80 SRM = 0.94 vs Lequesne Physical Function: ES = 0.83, SRM = 1.17		

Table 4

Reliability and Validity of the KOOS function, recreation and sport scale and KOOS-PS Short Form

Study	N	Population description	Internal consistency: Cronbach's alpha	Test-retest reliability: ICC	Comparator	Correlation
KOOS function, recreation and sport scale						
Xie et al 2006[22]	127	People from Singapore, China with knee OA scheduled for TKA	> 0.7	0.7	SF-36 Physical Functioning	r = 0.5
de Groot et al 2008[23]	262	People from the Netherlands with knee OA or TKA	> 0.9	> 0.5	SF-36 Physical Functioning	r = 0.1–0.6
Ornetti et al 2008[25]	67	People from France with knee OA or awaiting TKR	0.8	0.8	Osteoarthritis knee and hip quality of life questionnaire	r = 0.3
Goncalves et al 2009[24]	223	People from Portugal with symptomatic knee OA	0.9	0.9	SF-36 Physical Functioning	r = 0.6
Paradowski et al 2015[48]	68	People from Poland awaiting TKR	0.9	0.9	SF-36 Physical Functioning	r = 0.4
KOOS-PS Short Form						
Ornetti et al 2009[49]	87	People from France with knee OA		0.9	Osteoarthritis knee and hip quality of life questionnaire	r = −0.4
Goncalves et al 2010[26]	85	People from Portugal with symptomatic knee OA	0.9	0.9		
Davis et al 2009	248	People from Canada 6 months after TKR	0.9		WOMAC Physical Function	r = 0.9
Gul et al 2013[50]		People from Turkey with symptomatic knee OA	0.9	0.8	WOMAC Physical Function	r = 0.8
Singh et al 2014[51]	138	People from the US with knee OA		0.7		

Table 5

Sensitivity to change and Responsiveness of the KOOS function, recreation and sport scale and KOOS-PS Short Form

Study	N	ES and SRM of KOOS	MDC 95	MCID
KOOS Function, Sports, and Recreational Activity subscale				
Ornetti et al 2008[25]	67	ES= 1.3 SRM = 0.9		
Goncalves et al 2009[24]	223	ES = 0.8 SRM = 0.8		
Paradowski et al 2015[48]	68	ES = 1.6 SRM = 0.9	In individuals: 24.3 In groups: 2.9	
KOOS-PS Short Form				
Ornetti et al 2009[49]	87	ES = 0.5 SRM = 0.8		
Goncalves et al 2010[26]	85	ES = 0.9 SRM = 1.2		
Davis et al 2009[52]	248	SRM = 1.5 vs WOMAC Physical Function		
Singh et al 2014[51]	138	SRM = 1.5		MCID = 2.2 (0 to 100 scale)

Table 6

Reliability of PROMIS physical function

Study	N	Population description	Internal consistency: Cronbach's alpha	Test-retest reliability: ICC
Bartlett et al 2015[30]	177	People with Rheumatoid Arthritis at a routine clinic visit.	0.99	
Broderick et al 2013[53]	98	People with doctor diagnosed osteoarthritis		0.95
Paz et al 2013[54]	640	Adult Spanish-speaking Latinos	0.99	

Table 7

Construct validity of PROMIS Physical Function

Study	N	Comparator	Correlation
Bartlett et al 2015[30]	177	Modified Health Assessment Questionnaire	r = −0.8
Driban et al 2015[31]	204	SF-36 physical functioning	r = 0.8
Oude Voshaar et al 2015[29]	690	Health Assessment Questionnaire-Disability Index	r = −0.8