Scientific Foundation SPIROSKI, Skopje, Republic of Macedonia
Open Access Macedonian Journal of Medical Sciences. 2020 Jun 30; 8(F):133-136.
https://doi.org/10.3889/oamjms.2020.3914
eISSN: 1857-9655
Category: *F - Review Articles*
Section: *Narrative Review Article*

American Orthopedic Foot-and-Ankle Society Score, Where Are We Now? – A Narrative Review of Quality of Life Measures in Foot-and-Ankle Surgery

Jonathan James D'souza*

Warrington and Halton NHS Trust, Lovely Ln, Warrington WA5 1QG, United Kingdom

Edited by: Slavica Hristomanova-Mitkovska
Citation: D'souza JJ. American Orthopedic Foot-and-Ankle Society Score, Where Are We Now? - A Narrative Review of Quality of Life Measures in Foot-and-Ankle Surgery. Open Access Maced J Med Sci. 2020 Jun 30; 8(F):133-136. https://doi.org/10.3889/oamjms.2020.3914
Keywords: The American Orthopedic Foot and Ankle Society (AOFAS) score; Foot-and-ankle pathologies; Review article
***Correspondence:** Jonathan James D'souza, Warrington and Halton NHS Trust, Lovely Ln, Warrington WA5 1QG, United Kingdom. E-mail: Jonathan.dsouza@nhs.net
Received: 24-Oct-2019
Revised: 17-May-2020
Accepted: 20-May-2020

Funding: This research did not receive any financial support
Competing Interests: The authors have declared that no competing interests exist

Abstract

BACKGROUND: The American Orthopedic Foot and Ankle Society (AOFAS) score is the mostly widely used quality of life index for foot-and-ankle pathologies. However, there are many shortcomings of the QOL index with respect to validity, reliability, and responsiveness.

AIM: The narrative review will examine the AOFAS scoring system and review the outcome measure using the GRADE tool for quality of outcome measure.

CONCLUSIONS: Clinicians and researchers should cease using the AOFAS score as it is not a reliable tool. However, the recommendations suggested in this article can be used to upgrade the quality of the scoring system.

Introduction

There has been a paradigm shift in the assessment of primary outcome following treatment in trauma and orthopedics. The quality of evidence represents the degree to which a study design can support a given recommendation. Evidence refers to information obtained from current practice which can be randomized control trials, meta-analysis, case–control studies or expert opinions. Furthermore, the clinical data acquired from reviewing literature must be qualitatively appraised this would ensure that the reported information is appropriate to substantiate a precise recommendation. In the 90s, evidence-based working group suggested guidelines to systematically appraise the quality of data [1]. The levels of evidence were later used by the National Health Service in the United Kingdom to quantify clinical information and develop recommendations [2]. However, high levels of evidence if underpowered do not reflect a stronger recommendation [3].

In 2008, the WHO has recommended the grading of recommendations, assessment, development, and evaluations system to quantify strength of a clinical evidence. The GRADE system provides a structure for creating and submitting summaries of evidence and producing an organized method for developing clinical recommendations [4]. To begin with, the authors first decide the study population, the intervention and comparison and finally the outcome which has a marked effect on the conclusion. The next step is to assess the quality of data with respect to each outcome and subsequently rate the quality of evidence into high, moderate, low, and very low which is also referred to as – certainty of evidence. Since the GRADE tool is the most widely used quality of evidence measure and endorsed by majority of organization's around the world it will be used to assess quality of life measure in this essay. However, there are certain limitations to using the GRADE tool. First, GRADE cannot be used to assess studies that are based on risk of prognosis. Second, the application of the GRADE tool to well established clinical recommendations can be problematic. Third, GRADE proves to have limited application when applied to studies based on public health systems. Finally, GRADE has limited when applied to studies that evaluate clinical data on different interpretation of evidence [5].

The American Orthopedic Foot-and-Ankle Society (AOFAS) rating system and subscales are the most widely referenced foot-and-ankle scoring system used by researchers and clinical practitioners [6]. The system has four grading scales. The score is divided

into three subheadings – pain, function, and alignment. By combining the patient reported subjective scores of pain and function with surgeon's assessment of gait, sagittal-hindfoot motion and alignment a maximum score of 100 can be obtained [7]. There are several advantages to using the AOFAS system. First, it is a simple yet comprehensive method that can be used to report foot-and-ankle pathologies [7]. Second, the score has been scrutinized and has been reported as a reliable and valid outcome measurement tool [8]. These two qualities make it an effective tool that can be used to measure success of foot-and-ankle interventions.

There are many shortcomings of the AOFAS score. First, as the AOFAS score has a clinical examination component, differences in clinical examination skills among orthopedic surgeons potentially establishes a significant risk of inter-observer variability [9]. Subsequently, this has a cascading effect and affects the reliability of the rating system. In addition, the reliability of the score is further affected as there are no clear instructions whether the patient or the surgeon is required to complete the self-reported items on the questionnaire [10]. Second, by performing statistical analysis (Monte Carlo Modeling) of the score two conclusions was reported. Since, the score utilizes absolute descriptors, for example, "No limitation," "No difficulty" to describe answers on the score; it allows certain components of the score such as pain and function to be open to interpretation by the patient or clinician. Furthermore, a limited number of options available for each component of the score produces a skewed distribution as outcomes accumulate at the top and bottom end of the scale. The author concluded that AOFAS score was not a reliable tool [10]. Third, the absence of a logical theoretical structure on what the AOFAS score is supposed to measure affects the validity of the tool as an outcome measure for quality of life [11]. Fourth, for the AOFAS score to be an effective tool to measure quality of life it must be validated by examining its correlation to other validated QOL measures such as Short Form-36 (SF-36) and musculoskeletal function assessment score [12], [13]. Numerous studies have demonstrated that the AOFAS scores have poor correlation with other established outcome scores and therefore have inferior construct validity [6], [14], [15]. Finally, the point distribution for pain on the AOFAS scores (40 points) creates an inferior outcome measure in assessing foot-and-ankle pathologies with dominant symptoms of stiffness and deformity [15].

The GRADE tool rates the outcome of study and downgrades the quality based on risk of bias, imprecision, inconsistency, indirectness, and publication bias. The AOFAS score fails to define who the target population is, it only mentions that it is used to measure clinical outcomes in foot-and-ankle pathologies. Second, there is no "comparator," it merely explains the intervention in the form of the four subscales that can be used to measure foot-and-ankle pathology. The treatment outcome is reported as poor, fair, and good. Thus, the score fails to conceptualize a clear research question. The study also fails to mention how the participants were enrolled. This has a potential of introducing a high risk of bias of missing data and bias of selection of reported data [16]. Second, the study does not mention confidence intervals; therefore, the score is rated down for imprecision. Third, the score is downrated as several studies have shown inconsistencies in reliability and validity [6], [14], [15]. There are potentially four types of indirectness that can influence the quality of an outcome measure. Population indirectness can be introduced when there are differences between the study population and patients of interest. Indirectness of intervention occurs when there are differences between the intervention and the outcomes. Indirectness of comparison arises when there is no direct differentiation between two measured interventions. Finally, indirectness can be introduced when surrogate end points are measured instead of patient reported outcomes [17]. Publication bias can be introduced if the review does not include studies that are representative of the study. This can occur when including studies only published in one language are assessed in a review or selectively omitting studies as they do not reflect the intended results required to be published in a prominent journal.

A systematic review on QOL measures used in foot-and-ankle surgery by Button and Pinney [18] reviewed 256 articles from January 1990 to December 2001. The QOL measures would be rated based on validity, reliability, and responsiveness. Since the review compared QOL measures which were disease specific (foot function index), regional (AOFAS), and global (SF-36) this introduces an indirectness of comparison as there are no RCT's or observational studies which directly differentiate these interventions. Second, as some scores have been referenced and validated more than the others this creates a potential bias and consequently downgrades the outcome of one QOL measure over the other.

A recent systematic review by Safavi *et al*. [19] reviewed 669 publication and 76 foot-and-ankle scoring systems. The article defines the study population, intervention and comparisons. However, the study fails to explain what criteria were used to assess the quality of the selected studies. The review fails to mention what type (RCT's, observational) and number of studies included in the review. This introduces a bias and downgrades the quality of the review. The review concludes that foot pathologies such as rheumatoid arthritis were better measured with validated outcome scores (foot function index) while other foot-and-ankle deformities like ligament instability for ankles are better measured using the foot-and-ankle outcome score. However, the study fails to specify what qualitative assessment was carried to reach this conclusion. As

the review fails to include RCT's the quality of the study is further downgraded as the magnitude of effect is lessened and dose-response gradient is reduced [20]. Thus, the overall strength of recommendation of the study is weak as it has very low quality of evidence to support the summaries of evidence.

The recent American Foot-and-Ankle Society position statement has recommended that the score be discontinued as QOL measure since it has poor reliability, construct and criterion validity, and responsiveness [21]. Therefore, there is a need to select a more validated scoring system in foot and ankle to measure quality of life. This can be achieved by upgrading the AOFAS scoring system. First, there has be to a change in the points distribution for the pain sub score by reducing the allocated weightage of 40 points, this would make it more inclusive in measuring QOL outcomes in foot-and-ankle pathologies such as chronic ankle instability and hallux rigidus where dominant symptoms (stiffness, deformity) other than pain prevail. Second, absolute descriptors must be replaced in the QOL questionnaire this would reduce the ceiling effect of the scoring system and subsequently improving the reliability and validity of the scoring system [22]. Third, future systematic reviews on foot-and-ankle QOL measures must include randomized control trials this would, upgrade the quality of the study by improving magnitude of effect and dose-response gradient. The quality of evidence can also be upgraded by performing an extensive literature search and including studies reported in languages other than English this would reduce the publication bias. Finally, they must also provide for a method that allows standardization of the outcomes between QOL measures (foot and ankle) which would enable comparisons of treatment intervention thus reducing the indirectness of the study. Despite, being a non-validated outcome measure of QOL the AOFAS score is still the most widely utilized quality of life tool used in foot-and-ankle pathologies, by adopting the abovementioned points the score can be improved making it a superior QOL measure.

References

1. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents.Chest.1989;95(2):2S-4S. https://doi.org/10.1378/chest.95.2_supplement.2s
   PMid:2914516
2. Thornton J, Alderson P, Tan T, Turner C, Latchem S, Shaw E, *et al*. Introducing GRADE across the NICE clinical guideline program. J Clin Epidemiol. 2013;66(2):124-31.https://doi.org/10.1016/j.jclinepi.2011.12.007
   PMid:22406196
3. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg. 2011;128(1):305-10.
   PMid:21701348
4. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, *et al*. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.https://doi.org/10.1136/bmj.39489.470347.ad
   PMid:18436948
5. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, *et al*. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94.https://doi.org/10.1016/j.jclinepi.2010.04.026
   PMid:21195583
6. 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(4):494-8.https://doi.org/10.3113/fai.2007.0494
   PMid:17475146
7. 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(7):349-53.https://doi.org/10.1177/107110079401500701
   PMid:7951968
8. 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(2):65-74.https://doi.org/10.1053/j.jfas.2006.12.002
   PMid:17331864
9. Revell M, Pynsent P, Abudu A, Fairbank J. Trauma scores and trauma outcome measures. Trauma. 2003;5(1):61-70.https://doi.org/10.1191/1460408603ta268oa
10. Guyton GP. Theoretical limitations of the AOFAS scoring systems: An analysis using Monte Carlo modeling. Foot Ankle Int. 2001;22(10):779-87.https://doi.org/10.1177/107110070102201003
    PMid:11642529
11. Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty? Clin Orthop Relat Res. 2010;468(3):815-26. https://doi.org/10.1007/s11999-009-1036-y
    PMid:19672670
12. WareJE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473-83.https://doi.org/10.1097/00005650-199206000-00002
    PMid:1593914
13. 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(2):173-81.https://doi.org/10.1002/jor.1100140203
    PMid:8648493
14. Pena F, Agel J, Coetzee JC. Comparison of the MFA to the AOFAS outcome tool in a population undergoing total ankle replacement. Foot Ankle Int. 2007;28(7):788-93.https://doi.org/10.3113/fai.2006.0788
    PMid:17666170
15. 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(1):50-5.https://doi.org/10.1177/107110070302400108
    PMid:12540082
16. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, *et al*. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. BMJ.2016;355:i4919.
    PMid:27733354

17. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, *et al*. GRADE guidelines: 8. Rating the quality of evidence-indirectness. J Clin Epidemiol. 2011;64(12):1303-10. https://doi.org/10.1016/j.jclinepi.2011.04.014
PMid:21802903
18. 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(8):521-5.https://doi.org/10.1177/107110070402500802
PMid:15363371
19. Shazadeh Safavi P, Janney C, Jupiter D, Kunzler D, Bui R, Panchbhavi VK. A systematic review of the outcome evaluation tools for the foot and ankle. Foot Ankle Spec. 2019;12(5):461-70.https://doi.org/10.1177/1938640018803747
PMid:30338697
20. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, *et al*. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64(12):1311-6.https://doi.org/10.1016/j.jclinepi.2011.06.004
PMid:21802902
21. Pinsker E, Daniels TR. AOFAS position statement regarding the future of the AOFAS clinical rating systems. Foot Ankle Int. 2011;32(9):841-2.https://doi.org/10.3113/fai.2011.0841
PMid:22097157
22. Van Lieshout E, De Boer A, Meuffels D, Den Hoed P, Van der Vlies C, Tuinebreijer W, *et al*. American orthopaedic foot and ankle society (AOFAS) ankle-hindfoot score: A study protocol for the translation and validation of the Dutch language version. BMJ Open. 2017;7(2):e012884.https://doi.org/10.1136/bmjopen-2016-012884