

Measurement Properties of Total Ankle Replacement Outcome Measures: A Systematic Review

Ellie Pinsker, MD, Taucha M. Inrig, BScN, RN, MDiv, Timothy R. Daniels, MD, FRCSC, Dorcas Beaton, BScOT, MSc, PhD

Category: Ankle Arthritis

Keywords: Ankle arthroplasty, Psychometrics, Systematic review, Measurement properties, Best evidence synthesis

Introduction/Purpose: Researchers and clinicians measuring outcomes following total ankle replacement (TAR) are challenged by the wide range of outcome measures used in the literature without consensus as to which are valid, reliable, and responsive in this population. This review identifies region- or joint-specific outcome measures used for evaluating TAR outcomes and synthesizes evidence for their measurement properties.

Methods: A standard search strategy was conducted of electronic databases MEDLINE, EMBASE and CINAHL (to June 2015) to identify foot/ankle measures in use. A best evidence synthesis approach was taken to critically appraise measurement properties [COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)] of identified measures. The review was restricted to English publications and excluded cross-cultural adaptations. Measurement properties collected from each article were coded for validity, reliability, responsiveness, or interpretability. Psychometric evidence exists for identified measures tested in non-TAR populations, but were not the focus of this review.

Results: The search identified five measures to include in the best evidence synthesis, which have undergone psychometric testing in a TAR population (Foot Function Index, Ankle Osteoarthritis Scale, American Orthopaedic Foot and Ankle Society Ankle-Hindfoot Scale, Foot and Ankle Outcome Score, Self-Reported Foot and Ankle Score). Methodological quality of the studies providing psychometric evidence was assessed: (1) Validity—Two good quality studies examining different measures provide moderate evidence supporting construct validity (SEFAS; FFI, AOS, AOFAS *subjective items); (2) Reliability—two good quality studies examining different measures provide moderate evidence supporting internal consistency and test-retest reliability (FFI, AOS, AOFAS*; FAOS, SEFAS); (3) Responsiveness—one fair quality study provides limited evidence for AOFAS*, AOS, FFI; (4) Interpretability—two studies provided interpretability values (AOS, FFI, AOFAS*).

Conclusion: This review offers a basis for choosing the most appropriate instrument for evaluating TAR outcomes. Numerous outcome measures were identified with evidence supporting their use in populations with various foot/ankle conditions, but none have strong evidence supporting usage in a TAR population. Outcome measures must have adequate psychometric properties in all patient groups in which they are applied. Evidence supporting or critiquing an instrument should not be based on studies with poor quality methodology, as identified by this review. Further testing in a TAR population would benefit identified measures with emphasis on adequate sample sizes, testing a priori hypotheses, and evaluating their content validity for a TAR population.

Foot & Ankle Orthopaedics, 1(1)
DOI: 10.1177/ 2473011416500221
©The Author(s) 2016