Psychometric Properties of Commonly Used Outcomes Scores in Patients with Foot and Ankle Injuries

Foot & Ankle Category: Ankle

Author(s):
Karen K. Briggs, MPH, MBA
Lauren M. Matheny, BA
Thomas O. Clanton, MD
Diana C. Patterson, BA

Introduction
Many outcome scores are currently being used in patients with foot and ankle injuries. Although there have been several reviews of multiple scores, there is limited literature on all of these scores collected on the same patient. The purpose of this study was to determine the psychometric properties of commonly used foot and ankle outcome scores in a consecutive series of patients.

Methods
Six-hundred and seventy-two consecutive patients who presented with a foot or ankle injury completed a subjective questionnaire. From this questionnaire, the Foot and Ankle Disability Index (FADI) ADL, Sport and total score, the FAAM ADL, Sport, and total score, Tegner activity scale, Lysholm, SF12-Physical component, SF12-Mental component, WOMAC and the AOFAS hindfoot and AOFAS midfoot subjective component were calculated. The average age of the population was 47 years old (range: 18 to 85 years). There were 328 women and 344 men. One-hundred and eighty-six (28%) reported previous surgery and 21 were worker’s compensation cases.

Results
No score demonstrated normal distribution (p<0.05). For criterion validity, the foot and ankle scores were compared to the SF-12 Physical component. There was significant correlation with the Lysholm (rho=0.541; p<0.01), Tegner (rho=0.404; p<0.01), WOMAC (rho=-0.648; p<0.01), AOFAS hindfoot (rho=0.520; p<0.01), and AOFAS midfoot (rho=0.593; p<0.01). Weaker associations were seen with the FADI ADL (rho=0.163; p<0.01), FADI Sport (rho=0.143; p<0.01), FADI total (rho=0.170; p<0.01), FAAM ADL (rho=0.162; p<0.01) FAAM sport (rho=0.125; p<0.01) and FAAM total (rho=0.169; p<0.01). For content validity, no score had more than 5% of the scores at the lowest level (floor effect) or at the highest level (ceiling effect). To determine construct validity, 5 hypotheses were tested. (1) Patients with prior surgery would have lower scores; (2) Patients who rate their current level of ankle function as abnormal would have lower scores than those who consider their ankle nearly normal; (3) Scores would not correlate with patient age; (4) Patients with increased pain would have worse scores; (5) Scores would not differ by gender. The AOFAS hindfoot and Lysholm showed a significant difference for construct 1. All scores showed significant differences for construct 2. Tegner and WOMAC score failed construct 3 and were associated with age. All scores showed a significant correlation with pain (construct 4) except FADI sport and FAAM sport. All scores passed construct 5 except for Tegner which showed higher values in men.
Conclusion
This study shows that when all scores are compared in a single population, not all scores pass all validity tests. The Lysholm score is the only score that passed all three tests. All scores had adequate content validity (floor and ceiling effects) and criterion validity (as compared to SF-12). For construct validity (the 5 hypotheses), several scores were not able to distinguish between two different populations. More research is needed to determine which score is the best to measure outcomes in patients with foot and ankle injuries.