The Journal of Foot & Ankle Surgery 50 (2011) 420–429

Contents lists available at ScienceDirect

The Journal of Foot & Ankle Surgery

journal homepage: www.jfas.org

Validation of the American College of Foot and Ankle Surgeons Scoring Scales

Jeremy J. Cook, DPM, MPH, CPH, AACFAS [1], Emily A. Cook, DPM, MPH, CPH, AACFAS [1], Barry I. Rosenblum, DPM, FACFAS [2], Adam S. Landsman, DPM, PhD, FACFAS [3], Thomas S. Roukis, DPM, PhD, FACFAS [4]

[1] *Clinical Instructor in Surgery, Harvard Medical School, Beth Israel Deaconess Medical Center, Division of Podiatry, Boston, MA*
[2] *Assistant Clinical Professor of Surgery, Harvard Medical School, Beth Israel Deaconess Medical Center, Division of Podiatry, Boston, MA*
[3] *Assistant Professor of Surgery, Harvard Medical School and Chief, Division of Podiatric Surgery, Cambridge Health Alliance, Cambridge, MA; Assistant Professor, California School of Podiatric Medicine at Samuel Merritt University, Oakland, CA*
[4] *Attending Staff, Department of Orthopaedics, Podiatry, and Sports Medicine, Gundersen Lutheran Healthcare System, La Crosse, WI*

ARTICLE INFO

Level of Clinical Evidence: 2
Keywords:
foot function
health measurement
instrument
outcome
reliability
surgery
validation

ABSTRACT

The American College of Foot and Ankle Surgeons (ACFAS) assembled a task force to develop a scoring scale that could be used by the membership and practitioners-at-large. The original publication that introduced the scale focused primarily on use of the scale and provided only brief background on the development of the health measurement instrument. Concerns regarding the validity and reliability of the scale were raised within the professional community, and ACFAS assembled a task force to address these concerns. The purpose of this article is to address the issues raised by reporting the detailed methods used in the development of the ACFAS Scoring Scales. The authors who constitute this task force reviewed the body of work previously conducted and applied standards that serve to evaluate the scoring scale for: 1) validity, 2) reliability, and 3) sensitivity to change. The results showed that a systematic and comprehensive approach was used in the development of the scoring scales, and the task force concluded that the statistical methods and instrument development process for all 4 modules of the scoring scales were conducted in an appropriate manner. Furthermore, modules 1 and 2 have been rigorously assessed and the elements of these modules have been shown to meet standards for validity, reliability, and sensitivity to change.

The American College of Foot and Ankle Surgeons (ACFAS) Universal Evaluation Scoring Scale Task Force developed 4 anatomically based scoring scales intended as clinical instruments to be used to measure subjective and objective parameters germane to foot and ankle surgery. Modules 1 and 2 were released in 2002 (1, 2), and a subsequent user's guide containing all of the modules was published in 2005 (3). The first 2 modules, which focus on the first metatarsophalangeal joint and first ray, and the forefoot, respectively, were developed and statistically analyzed in an effort to confirm that their development and design yield valid results.

The work of the original task force a priori was intended to be periodically reevaluated after a reasonable trial period by practicing surgeons. However, reevaluation has not yet occurred and the scoring scales have not yet been widely used for a variety of reasons, including issues raised about the accuracy of validation by external (4) and internal reviews by the ACFAS Evidence Based Medicine and Research Committee. Specifically, the methods, analysis, and results of the original development process were not reported and have subsequently been challenged (4).

The purpose of this report is to address the issues raised by reporting the detailed methods used in the development of the ACFAS Scoring Scales. In addition, the steps used for prior validation of modules 1 and 2, including assessments of validity, reliability, and sensitivity to change, are presented here as well.

Materials and Methods

The authors, who comprise the current ACFAS Scoring Scale Task Force, were assembled in 2010 to conduct the first reevaluation and to address the limitations of previous publications related to the ACFAS Scoring Scales. This process involved collection of all materials archived by the ACFAS that pertained to the development of the original scoring scales. We then reviewed the previous work to assess the conclusions that were reported in the original publications. Specific objectives of the current task force were derived from our own independent review of the available materials, and the criticisms raised by others and mentioned earlier in this report. In addition, we report the methods by which the original scoring scales were developed and the approach taken for instrument validation. Where possible, we determined standardization of the specific radiographic and objective functional items used in modules 1 and 2 to maximize reproducibility. To achieve this, we performed a detailed

Financial Disclosure: None reported.
Conflict of Interest: None reported.

Address correspondence to: Thomas S. Roukis, DPM, PhD, FACFAS, Attending Staff, Department of Orthopaedics, Podiatry, and Sports Medicine, Gundersen Lutheran Healthcare System, 2nd Floor Founders Building, 1900 South Avenue, La Crosse, WI 54601.
E-mail address: tsroukis@gundluth.org (T.S. Roukis).

1067-2516/$ - see front matter © 2011 by the American College of Foot and Ankle Surgeons. All rights reserved.
doi:10.1053/j.jfas.2011.03.005

electronic search of the literature to determine if the techniques described possess a higher level of inter-rater and intra-rater reliability than those originally selected. Our objective was to strengthen the current scoring scale by identifying widely recognized and accepted techniques for making measurements, and bringing a higher level of standardization to the scoring process. The genesis of the scoring scales is depicted schematically in Figure 1.

Results

The 4 modules of the ACFAS Universal Evaluation Scoring Scales are depicted in Figures 2 through 5, and the user instructions that accompany the first 2 modules are depicted in Figure 6. The scoring scales, as well as the work completed by the original task force, served as the starting point for our analysis of the development of the instrument in terms of validity, reliability, and sensitivity to change. The methods of radiographic and functional measurement described in Figure 6 represent techniques that are generally known to show high levels of inter-rater and intra-rater reliability.

Instrument Development

After a detailed review of all correspondence and materials archived by the ACFAS for the scoring scales, we determined that the original task force initiated its scoring scale development with a review of the literature including previously published scoring scales. A survey was conducted at the 1999 ACFAS Annual Scientific Conference asking members to rank the importance of having a scoring scale and the importance of different parameters. Of those members surveyed, 88% agreed or strongly agreed that the development of a foot and ankle surgery scoring scale was a worthwhile project for the ACFAS to conduct. The parameters were ranked by the members from the most to the least important, as follows: pain, function, shoe gear limitations, radiographic measurements, and cosmesis. The original task force then consulted additional experts in foot and ankle surgery to create initial scoring scales (i.e., sections, subsections, questions and answers, weighting of answers). This evolved into a modified Delphi process, a standard procedure in which a panel of experts is assembled, to obtain consensus for the conceptual framework of the specific questions and answers generated (5). Subsequently, the format of these questions and answers was revised through the modified Delphi process. Patient focus group sessions were then conducted, and the findings were used to incorporate patients' values and interpretations of the questions into the instruments (Table 1). Global questions were also derived to address the pain, appearance, and functional components of the scoring scales (Figure 1).

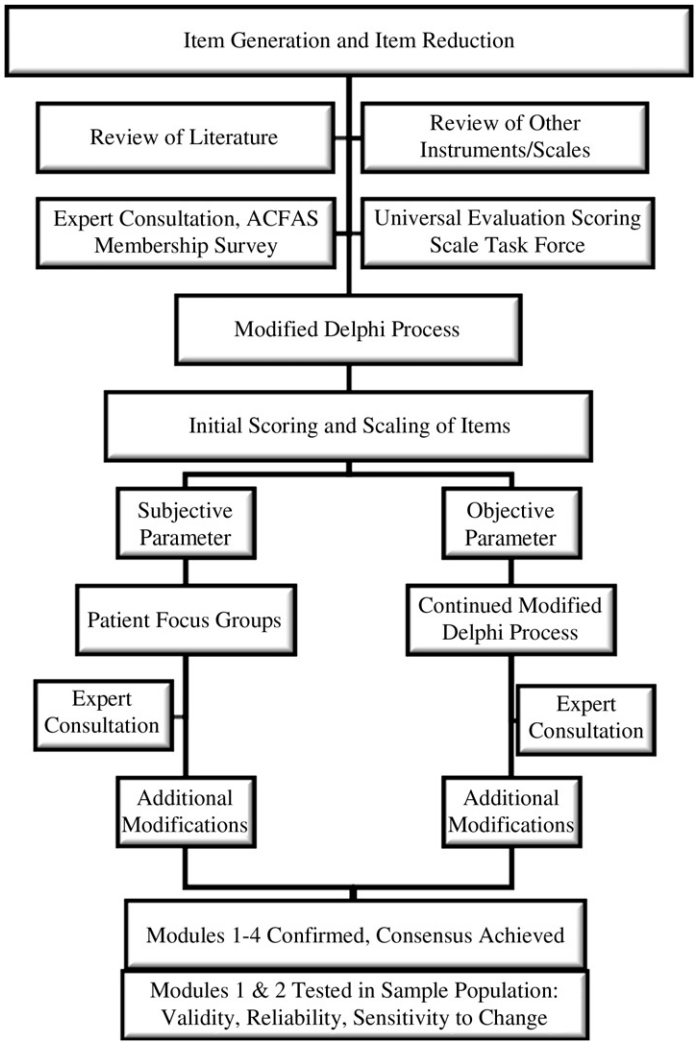

Fig. 1. ACFAS Scoring Scales genesis.

Final ACFAS Scoring Scale

The final ACFAS Scoring Scales (Figures 2 through 5) are comprised of the following 4 modules:

Module 1: First Metatarsophalangeal Joint (MPJ) and First Ray (11 questions)
Module 2: Forefoot (Excluding First Ray) (12 questions)
Module 3: Rearfoot (Including Flatfoot) (16 questions)
Module 4: Ankle (22 questions)

Each of the final scoring scales included a total of 100 points (50 subjective, 50 objective). The original task force selected a total of 100 points for ease of use, interpretability, and the ability to maintain analogy with other scoring scales. The subjective parameters consisted of subsections encompassing questions and answers pertaining to pain (30 points), appearance (cosmesis) (5 points), and functional capacity (15 points), whereas the objective parameters consisted of radiographic (18 points) and functional (musculoskeletal) (32 points) measurements. Item weighting was determined by means of a process that included expert consultation, the modified Delphi process, and patient focus groups.

The primary aim of the ACFAS Scoring Scale is to evaluate the subjective and objective health outcomes before and after foot and ankle surgery. The ACFAS Scoring Scale should be used with patients enrolled in prospective clinical trials for foot and ankle surgery under the following conditions:

- Pathology/disease: Foot and ankle musculoskeletal diseases requiring surgical intervention
- Population for intended use: Adults ($\geq$18 years old), English speaking
- Administration mode: Subjective component, self-administered; objective component, clinician-rated
- Recall/observation period: Condition at present time administered preoperatively and postoperatively, except subjective pain response is over the past month

The ACFAS Scoring Scales (Figures 2 through 5) modules 1 and 2 were tested in a total of 91 patients in module 1 and 84 patients in module 2, in 6 centers over several years for validity, reliability, and sensitivity to change.

ACFAS Universal Evaluation Scoring Scale
Module 1: First Metatarsophalangeal Joint (MPJ) & First Ray

Patient: _____
Date: _____
Patient #: _____

Instructions to the Patient: Please answer the following questions honestly with regard to the condition of your foot.

Patient Subjective Questionnaire

1. Pain (30 points)
   Over the past month, how much has your foot pain limited your daily activities?
   - ☐ I have no pain with normal activities (30)
   - ☐ I have slight or occasional pain but no limitation of activities (22)
   - ☐ I have moderate pain limiting some activities (14)
   - ☐ I have pain and significant limitation of activities (6)
   - ☐ I have severe pain that limits almost all activity (0)

2. Appearance (5 points)
   How would you rate the appearance of your big toe joint?
   - ☐ I like it very much (5)
   - ☐ I mostly like it (4)
   - ☐ I'm not sure either way – neutral (3)
   - ☐ I mostly do not like it (2)
   - ☐ I definitely dislike it (0)

3. Functional Capacities (15 points)
   How frequently do you have pain while wearing shoes?
   - ☐ I am able to continuously wear any type of shoe (15)
   - ☐ I am able to war any type of shoe most of the time (10)
   - ☐ I am able to war *only* walking, athletic or casual shoes (5)
   - ☐ I am able to wear only special order, orthopedic or custom-made shoes (0)

Objective Parameters

4. Radiographic Evaluation (18 points)
   AP (weight bearing dorsoplantar) View (6 points)
   HA (Hallux Abductus) Angle (6 points)
   - ☐ 0º-20º (6)
   - ☐ 21º-30º (3)
   - ☐ ≥ 31º (0)
   - ☐ -1º to -3º (2)
   - ☐ > -3º (0)

   IM (Intermetatarsal) Angle (6 points)
   - ☐ 0º-10º (6)
   - ☐ 11º-19º (3)
   - ☐ ≥ 20º (0)
   - ☐ < 0º (0)

   Lateral View (6 points)
   First Metatarsal Declination Angle (6 points)
   - ☐ 16º-24º (6)
   - ☐ 25º-29º (3)
   - ☐ ≥ 29º (0)
   - ☐ 10º-15º (2)
   - ☐ > 10º (0)

5. Function (32 points)
   Hallux Purchase (Paper-pull out test) (10 points)
   - ☐ Not Movable (10)
   - ☐ Resistant (5)
   - ☐ Easy (0)

   Range of Motion: First Ray (17 points)
   First MPJ Dorsiflexion (11 points)
   - ☐ ≥ 60º (11)
   - ☐ 45º-59º (8)
   - ☐ 36º-44º (4)
   - ☐ < 36º (0)

   First MPJ Plantarflexion (4 points)
   - ☐ ≥ 0º (4)
   - ☐ < 0º (0)

   Hallux IPJ Extension (2 points)
   - ☐ Extend to 0º (2)
   - ☐ < 0º (0)

   Limp from Foot Pain (without shoes) (5 points)
   - ☐ No (5)
   - ☐ Yes (0)

Subjective Points: _____
Objective Points: _____
Total Points Module 1: _____

Fig. 2. Module 1.

ACFAS Universal Evaluation Scoring Scale
Module 2: Forefoot (Excluding First Ray)

Patient: _____
Date: _____
Patient #: _____

Patient Questionnaire

1. Pain (30 points)
   Over the past month, how does your foot pain limit your daily activities?
   - ☐ I have no pain with normal activities (30)
   - ☐ I have slight or occasional pain but no limitation of activities (22)
   - ☐ I have moderate pain limiting some activities (14)
   - ☐ I have pain and significant limitation of activities (6)
   - ☐ I have severe pain that limits almost all activity (0)

2. Appearance (Cosmesis) (5 points)
   How do you rate the appearance of your foot?
   - ☐ I like it very much (5)
   - ☐ I mostly like it (4)
   - ☐ I'm not sure either way – neutral (3)
   - ☐ I mostly do not like it (2)
   - ☐ I definitely dislike it (0)

3. Functional Capacities (15 points)
   How frequently do you have pain while wearing shoes?
   - ☐ I am able to continuously wear any type of shoe (15)
   - ☐ I am able to wear any type of shoe most of the time (10)
   - ☐ I am able to wear *only* walking, athletic, or casual shoes (5)
   - ☐ I am able to wear *only* special order, orthopedic, or custom shoes (0)

Objective Parameters

4. Radiographic (18 points)
   (If studying more than one ray, evaluate each ray separately)
   4-5 Intermetatarsal Angle (4 points) (Refs. 3, 4)
   - ☐ 0º-8º (4)
   - ☐ ≥ 9º (0)

   Metatarsal Length (10)
   - ☐ 3-6 mm change from pretreatment length (10)
   - ☐ ≤ 2 mm change from pretreatment length (0)
   - ☐ ≥ 7 mm change from pretreatment length (0)

   Transverse Plane Position (4)

   MPJ
   - ☐ 0º-5º ab/adduction (2)
   - ☐ 5º (0) (Ref. 5)

   IPJ
   - ☐ 0º-5º ab/adduction (2)
   - ☐ 5º (0)

5. Function (32 points) (If studying more than one ray, evaluate each ray separately)
   Lesser MPJ Range of Motion (15)
   Dorsiflexion (7)
   - ☐ ≥ 65º (7)
   - ☐ 45º-64º (3)
   - ☐ ≤ 45º (0)

   Plantarflexion (8)
   - ☐ ≥ 0º (8)
   - ☐ < 0º (0)

   Digital Purchase (4)
   - ☐ Yes (4)
   - ☐ No (0)

   Drawer Sign/Dislocation (8)
   - ☐ Stable (8)
   - ☐ Subluxable (4)
   - ☐ Dislocated (0)

   Limp from Foot Pain (without shoes) (5)
   - ☐ No (5)
   - ☐ Yes (0)

Subjective Points: _____
Objective Points: _____
Total Points Module 2: _____

Fig. 3. Module 2.

ACFAS Universal Evaluation Scoring Scale
Module 3: Rearfoot (Including Flatfoot)

Patient: _____
Date: _____
Patient #: _____

Instructions to the Patient: Please answer the following questions honestly with regard to the condition of your foot.

Patient Subjective Questionnaire

1. Pain (30 points)
   Over the past month, how much has your foot pain limited your daily activities?
   - ☐ I have no pain with normal activities (30)
   - ☐ I have slight or occasional pain but no limitation of activities (22)
   - ☐ I have moderate pain limiting some activities (14)
   - ☐ I have pain and significant limitation of activities (6)
   - ☐ I have severe pain that limits almost all activity (0)

2. Appearance (5 points)
   How would you rate the appearance of your big toe joint?
   - ☐ I like it very much (5)
   - ☐ I mostly like it (4)
   - ☐ I'm not sure either way – neutral (3)
   - ☐ I mostly do not like it (2)
   - ☐ I definitely dislike it (0)

3. Functional Capacities (15 points)
   Ability to function on one flight of stairs, incline, & uneven terrain (7 points)
   - ☐ I have no difficulty (7)
   - ☐ I have some difficulty (4)
   - ☐ I have significant difficulty (0)

   Walking Aids (5 points)
   - ☐ I do not use any (5)
   - ☐ I wear a prescription brace above my ankle (3)
   - ☐ I use one cane or one crutch (3)
   - ☐ I use either crutches, a walker or a wheelchair (0)

   Shoes (3 points)
   - ☐ I can wear normal shoes (3)
   - ☐ I am able to wear *only* walking or athletic shoes (2)
   - ☐ I am able to wear *only* special order, orthopedic or custom-made shoes (0)

Objective Parameters

4. Radiographic Evaluation (18 points)
   Lateral (Weight bearing) View
   Sagittal Plane Talo-First Metatarsal Declination Angle (3 points)
   - ☐ 0º ± 5º (3)
   - ☐ < 5º (0)

   Calcaneal Inclination Angle (3 points)
   - ☐ 15º-25º (3)
   - ☐ < 15º or >25º (0)

   Frontal Plant (Weight Bearing Long Leg Calcaneal Axial) View
   Calcaneal-Tibial Angle (6 points)
   - ☐ ≤ 2º varus to ≤ 4º valgus (6)
   - ☐ > 2º varus to > 4º valgus (0)

   Calcaneal Translational Displacement (6 points)
   - ☐ 5-10 mm lateral (6)
   - ☐ < or > 5-10 mm lateral (0)

   AP Weight Bearing View (Transverse Plane Measurements)
   Talo-First Metatarsal Declination Angle (3 points)
   - ☐ 0 ± 5º (3)
   - ☐ > 5º (0)

   Talo-calcaneal Angle (6 points)
   - ☐ 15º-30º (3)
   - ☐ < 15º or > 30º (0)

5. Function (32 points)
   Range of Motion (18 points)
   Ankle Dorsiflexion – knee extended (7 points)
   - ☐ 10º-15º (7)
   - ☐ 16º-20º or 5º-9º (4)
   - ☐ > 20º or < 5º (0)

   Ankle Plantarflexion (4 points)
   - ☐ ≥ 30º (4)
   - ☐ 15º-29º (2)
   - ☐ < 15º (0)

   Rearfoot Subtalar Joint Inversion/Eversion (7 points)
   - ☐ 25º-45º (7)
   - ☐ 10º-24º (4)
   - ☐ < 10º or > 24º (0)

   Single Limb Heel Rise (9 points)
   - ☐ Able to perform multiple rises without difficulty (9)
   - ☐ Performed with some difficulty (4)
   - ☐ Unable to perform (0)

   Limp due to Foot Pain (without shoes) (5 points)
   - ☐ No (5)
   - ☐ Yes (0)

Subjective Points: _____

Objective Points: _____
Total Points Module 3: _____

Fig. 4. Module 3.

ACFAS Universal Evaluation Scoring Scale
Module 4: Ankle

Patient: _____
Date: _____
Patient #: _____

Instructions to the Patient: Please answer the following questions honestly with regard to the condition of your foot.

Patient Questionnaire

1. Pain (30 points)
   - Over the past month, how much has your foot pain limited your daily activities?
     - ☐ I have no pain with normal activities (30)
     - ☐ I have slight or occasional pain but no limitation of activities (22)
     - ☐ I have moderate pain limiting some activities (14)
     - ☐ I have pain and significant limitation of activities (6)
     - ☐ I have severe pain that limits almost all activity (0)

2. Appearance (5 points)
   - How would you rate the appearance of your big toe joint?
     - ☐ I like it very much (5)
     - ☐ I mostly like it (4)
     - ☐ I'm not sure either way – neutral (3)
     - ☐ I mostly do not like it (2)
     - ☐ I definitely dislike it (0)

3. Functional Capacities (15 points)
   - Ability to function on one flight of stairs, incline, & uneven terrain (7 points)
     - ☐ I have no difficulty (7)
     - ☐ I have some difficulty (4)
     - ☐ I have significant difficulty (0)
   - Walking Aids (5 points)
     - ☐ I do not use any (5)
     - ☐ I wear a prescription brace above my ankle (3)
     - ☐ I use one cane or one crutch (3)
     - ☐ I use either crutches, a walker or a wheelchair (0)
   - Shoes (3 points)
     - ☐ I can wear normal shoes (3)
     - ☐ I am able to wear *only* walking or athletic shoes (2)
     - ☐ I am able to wear *only* special order, orthopedic or custom-made shoes (0)

Objective Parameters

4. Radiographic Evaluation (18 points)
   - Anteroposterior (AP) View (6 points)
     - Lateral Distal Tibial Angle (LDTA), anatomic axis (3 points)
       - ☐ LDTA between 86º to 92º (3)
       - ☐ LDTA < 86º or > 92º (0)
     - Tibial Talar Angle (3 points)
       - ☐ 0º-4º varus or valgus (3)
       - ☐ 5º-9º varus or valgus (2)
       - ☐ ≥ 10º varus or valgus (0)
   - Mortise View (3)
     - Talocrural Angle (3 points)
       - ☐ 83º ± 4º (3)
       - ☐ < 79º or > 87º (0)
       - ☐ > 5º variance compared to contralateral ankle (0)
   - Long Leg Calcaneal Axial Weight Bearing View (3 points)
     - Calcaneal-Tibial Angle (3 points)
       - ☐ ≤ 2º varus to ≤ 4º valgus (3)
       - ☐ > 2º varus to > 4º valgus (0)
   - Lateral View (6 points)
     - Anterior Distal Tibial Angle (ADTA), anatomic axis (3 points)
       - ☐ ADTA between 78º-82º (3)
       - ☐ ADTA < 78º or > 82º (0)
     - Center of Rotational Axis of the Ankle (3 points)
       - ☐ Talar lateral process directly under anatomic axis of tibia (3)
       - ☐ Displacement > 1 cm anterior or posterior (0)
   - Optional Section: Special Considerations (Deductions)
     - Calcaneal Translational Displacement (-3 points)
       - ☐ < 5 mm or > 10 mm lateral (-3)
     - Joint Space Thickness (-3 points)
       - ☐ < 50% reduction (-3)
     - Tibial Fibular Overlap (-3 points)
       - ☐ < 6 mm (-3)
     - Medial Clear Space (-3 points)
       - ☐ ≥ 4 mm (-3)
     - Stress Inversion (-3 points)
       - ☐ ≥ 5º compared to contralateral ankle (-3)
     - Stress Anterior Drawer (-3 points)
       - ☐ Positive (> 4 mm) (-3)

5. Function (32 points)
   - Range of Motion (18 points)
     - Ankle Dorsiflexion – knee extended (7 points)
       - ☐ 10º-15º (7)
       - ☐ 16º-20º or 5º-9º (4)
       - ☐ > 20º or < 5º (0)
     - Ankle Plantarflexion (4 points)
       - ☐ ≥ 30º (4)
       - ☐ 15º-29º (2)
       - ☐ < 15º (0)
     - Rearfoot (Calcaneal) Subtalar Joint Inversion/Eversion (7 points)
       - ☐ 25º-45º (7)
       - ☐ 10º-24º (4)
       - ☐ < 10º or > 45º (0)
   - Balance (Rhomberg Test) (9 points)
     - ☐ 15-20 seconds (9)
     - ☐ 5-14 seconds (4)
     - ☐ < 5 seconds (0)
   - Limp due to Ankle Pain (without shoes) (5 points)
     - ☐ No (5)
     - ☐ Yes (0)

Subjective Points: _____
Objective Points: _____
Total Points Module 4: _____

Fig. 5. Module 4.

Module	Measurement	Methodology and reference citation
1	Hallux abductus (HA) angle	Marks are placed in the mid-diaphyseal region of the proximal phalanx and the first metatarsal at an equal distance from the medial and lateral cortices. The hallux valgus angle is formed by intersection of the diaphyseal axes of the proximal phalanx and the first metatarsal. (Coughlin MJ, Freund E. The reliability of angular measurements in hallux valgus deformities. Foot Ankle Int. 22(5):369-379, 2001.)
	Intermetatarsal (IM) angle (1st and 2nd metatarsals)	Mid-diaphyseal points are marked equidistant from the medial and lateral cortices of the first and second metatarsals in the proximal and distal mid-diaphyseal region and the longitudinal axis is drawn for both the first and second metatarsals. The intermetatarsal angle is formed by the intersection of the axes of the first metatarsal and the second metatarsal. (Coughlin MJ, Freund E. The reliability of angular measurements in hallux valgus deformities. Foot Ankle Int. 22(5):369-379, 2001.)
	First metatarsal declination angle	The first metatarsal declination angle is mean angle between longitudinal axis of the first metatarsal bone and the ground. The longitudinal axis of the first metatarsal declination angle was defined as centerline between the dorsal and plantar cortical bone. The first metatarsal declination angle is measured by obtaining a bisection of the head and base of the first metatarsal and measuring this line to the ground plane. (Bouaicha S, Ehrmann C, Moor BK, Maquieira GJ, Espinosa N. Radiographic analysis of metatarsus primus elevatus and hallux rigidus. Foot Ankle Int. 31(9):807-814, 2010.)
	Hallux purchase (paper-pull out test)	The strength of the plantarflexor muscles of the hallux is determined using the paper-pull out or grip test. The subject should be seated with the hip, knee, and ankle at 90 degrees and instructed to use their hallux toe muscles to push down on a standard business card while the examiner stabilizes their ankle and forefoot and attempted to slide the cardboard away from the toe. The duration of the test is 3 to 4 seconds. The test is documented as not movable (pass) if the subject can hold the card for the full time period or easy (fail) if the subject failed to grip the card at all. No criteria have been formally studied regarding the resistant choice available. (Menz HB, Zammit GV, Munteanu SE, Scott G. Plantarflexion strength of the toes: Age and gender differences and evaluation of a clinical screening test. Foot Ankle Int. 27(12):1103-1108, 2006.)
	First metatarsophalangeal joint (MPJ) dorsiflexion	The static non-weight bearing dorsiflexion of the first metatarsophalangeal joint is performed with the subject sitting lengthwise on a treatment table with the foot positioned over the edge of the table. The first metatarsophalangeal joint is identified by dorsiflexion and plantarflexion of the hallux. A mark is made at the medial aspect of the center of the first metatarsophalangeal joint. After palpation, lines bisecting the medial aspect of the shafts of both the first metatarsal and proximal phalanx are drawn on the medial side of the foot. These two lines are then connected to the joint center marking. Assisted non-weight bearing dorsiflexion of the first metatarsophalangeal joint is determined with the subtalar joint in neutral position and midtarsal joint stabilized by applying force to the fourth and fifth metatarsal heads in the direction of dorsiflexion. The investigator maintains this force on the fourth and fifth metatarsal heads while aligning the arms of the goniometer on the first metatarsophalangeal joint reference lines. The investigator exerts pressure on the proximal phalanx of the subjects' hallux in a dorsal direction utilizing the bisection lines above and the angle measured with a goniometer. The value recorded is the maximum angle at which the hallux could not be passively moved into further dorsiflexion. (Hopson MM, McPoil TG, Cornwall MW. Motion of the first metatarsophalangeal joint. Reliability and validity of four measurement techniques. J Amer Podiatr Med Assoc 85(4):198-204, 1995.)
	First metatarsophalangeal joint (MPJ) plantarflexion	The static non-weight bearing plantarflexion of the first metatarsophalangeal joint is performed with the subject sitting lengthwise on a treatment table with the foot positioned over the edge of the table. The first metatarsophalangeal joint is identified by dorsiflexion and plantarflexion of the hallux. A mark is made at the medial aspect of the center of the first metatarsophalangeal joint. After palpation, lines bisecting the medial aspect of the shafts of both the first metatarsal and proximal phalanx are drawn on the medial side of the foot. These two lines are then connected to the joint center marking. Assisted non-weight bearing plantarflexion of the first metatarsophalangeal joint is determined by the investigator exerting pressure on the proximal phalanx of the subjects' hallux in a plantar direction utilizing the bisection lines above and the angle measured with a goniometer. The value recorded is the maximum angle at which the hallux could not be passively moved into further plantarflexion. (Resch S, Ryd L, Stenström A, Johnson K, Reynisson K. Measuring hallux valgus: a comparison of conventional radiography and clinical parameters with regard to measurement accuracy. Foot Ankle Int. 16(5):267-270, 1995.)
2	4-5 intermetatarsal angle	Mid-diaphyseal points are marked equidistant from the medial and lateral margins of condyles of the fourth and fifth metatarsals and the longitudinal axis is drawn for both the fourth and fifth metatarsals. The 4-5 intermetatarsal angle is formed by the intersection of the axes of the fourth metatarsal and the fifth metatarsal. (Karasik D. Preoperative assessment of symptomatic bunionette deformity: Radiographic findings. Am J Radiol 164:147-149, 1995.)
	Metatarsal length	The axis of the forefoot is first identified between the midpoint of the distance between the medial aspect of the talar head and the distal lateral edge of the calcaneus and the tip of the second metatarsal head. The distance in millimeters between lines starting at the tips of the first and second metatarsal heads, perpendicular to the axis of the forefoot is then performed. Similar distances are measured between the second and third, third and fourth, and fourth and fifth metatarsal heads. The value is denoted as positive if the medial metatarsal is longer than the next adjacent lateral one (i.e., first metatarsal longer than the second metatarsal) and negative if the medial metatarsal is shorter than the next adjacent one (i.e., second metatarsal shorter than the third metatarsal). (Deleu PA, Pod H, Leemrijse T, Birch I, Vande Berg B, Bevernage B. Reliability of the Maestro radiographic measuring tool. Foot Ankle Int 31(10):884-891, 2010.)
	Transverse plane position, metatarsophalangeal joint (MPJ)	The angle of deviation of the lesser metatarsophalangeal joints is obtained by using two mid-diaphyseal reference points on each respective proximal phalanx and metatarsal. These points are placed 1 to 2 cm from the articular surfaces and are connected by a longitudinal line that forms the longitudinal axis of the specific bone. When the digit is medially deviated (adducted) at the corresponding metatarsophalangeal joint, the value denoted as positive, and when the digit is laterally deviated (abducted) the value denoted as negative. (Kaz AJ, Coughlin MJ. Crossover second toe: Demographics, etiology, and radiographic assessment. Foot Ankle Int 28(12):1223-1237, 2007.)
	Transverse plane position, interphalangeal joint (IPJ)	The angle of deviation of the lesser toes interphalangeal joints is obtained by using the mid-diaphyseal reference points on each respective proximal and intermediate or distal and intermediate phalanx. A longitudinal line that forms the longitudinal axis of the specific bone connects these points. When the digit is medially deviated (adducted) at the corresponding interphalangeal joint, the value denoted as positive, and when the digit is laterally deviated (abducted), the value denoted as negative. (Coughlin MJ. Lesser toe abnormalities. Instr Course Lect 52: 421-444, 2003.)
	Digital purchase (paper-pull out test)	The strength of the plantarflexor muscles of the lesser toes is determined using the paper-pull out or grip test. The subject should be seated with the hip, knee, and ankle at 90 degrees and instructed to use their lesser toe muscles to push down on a standard business card while the examiner stabilizes their ankle and forefoot and attempted to slide the cardboard away from the toe. The duration of the test is 3 to 4 seconds. The test is documented as not movable (pass) if the subject can hold the card for the full time period or easy (fail) if the subject failed to grip the card at all. No criteria have been formally studied regarding the resistant choice available. (Menz HB, Zammit GV, Munteanu SE, Scott G. Plantarflexion strength of the toes: Age and gender differences and evaluation of a clinical screening test. Foot Ankle Int. 27(12):1103-1108, 2006.)
	Drawer sign/dislocation	The investigator grasps the proximal phalanx between their thumb and forefinger of one hand the corresponding metatarsal is grasps between the thumb and forefinger of their other hand. The toe is distracted in a dorsal direction. When plantar pain is elicited metatarsophalangeal joint instability is present. If the lesser toe rests in a fixed position dorsal to the metatarsal head it is considered dislocated. If the lesser toe cannot be distracted dorsally it is considered stable. The toe is considered subluxable when it is neither stable nor dislocated. (Coughlin MJ. Second metatarsophalangeal joint instability in the athlete. Foot Ankle 14:309-319, 1993.)

Fig. 6. Instructions for the objective section of the modules 1 and 2 of the ACFAS Universal Evaluation Scoring Scales.

Validity

After a detailed review as noted above, we determined that the first validation study assessed both content (face) and initial construct validity with an a priori plan to assess criterion validity and more in-depth assessment of construct validity in future scoring scale updates. Content (face) validity was thoroughly assessed through the modified Delphi process consisting of 6 members and 2 consultants (3) in

Table 1
Patient focus groups: Reasons for having foot surgery from a patient perspective

	Module 1			Module 2		
	Appearance	Pain	Inability to Wear All Shoe Types	Appearance	Pain	Inability to Wear All Shoe Types
Presurgery/test	5.5%	86.8%	7.7%	7.2%	84.5%	8.3%
Presurgery/retest	8.7%	79.7%	11.6%	12.3%	78.5%	9.2%
Kappa coefficient			0.44[3]			0.40[3]

0 = poor agreement; 1 = slight agreement; 2 = fair agreement; 3 = moderate agreement; 4 = substantial agreement; 5 = almost perfect agreement (17).

collaboration with an independent biostatistician over several years' time. This helped ensure adequate content (face) validity and appropriateness of questions and answers relative to the purpose of this scoring scale. A review of the content by the current task force reaffirmed the appropriateness of questions and answers, and helped to identify and refine areas where questions could be improved through further standardization.

Construct validity, as it relates to the proposed use of the scales, refers to the anticipated behavior of the scale after surgical intervention. Construct validity was demonstrated via the expected directional change in the module scores after surgical intervention (e.g., after first metatarsal osteotomy for hallux valgus, it is expected that the intermetatarsal 1–2 angle would decrease rather than increase). We hypothesized that the scoring scale would increase between the specific preoperative state and postoperative surgical intervention, because higher scores represent more desirable outcomes. A paired 2-sided Student's *t* test using the preoperative and postoperative scores found a consistent increase in the total scores with the expected directionality ($P < .01$) (Tables 2 and 3).

Reliability

After a detailed review as noted above, we have identified the following process as occurring. Reliability of the subjective portion of the scoring scales was assessed via test–retest with an a priori plan to assess internal consistency in future scoring scale updates. Test–retest analysis was conducted with correlation coefficients between the initial test and the retest 7 to 10 days later. In broad terms, test–retest is a technique used to confirm reliability by answering and then re-answering the same question at a later time. Initial and subsequent answers are compared to determine whether the question can be reliably answered. We confirmed that a test–retest evaluation was performed both preoperatively and postoperatively such that each question and answer were collected 4 times. Fair to substantial agreement was obtained in all categories (Tables 4 and 5).

Sensitivity to Change

Sensitivity is based on the ability of the scoring scales to reflect a change after an intervention where a change would be reasonably expected (i.e., after surgery). After a detailed review, as described above, we have identified the following process as having occurred. The ACFAS Scoring Scale was administered preoperatively and again within a 6-month postoperative period. A paired 2-sided Student's *t* test was used to evaluate the mean change in the total score. A statistically significant change in the total score was detected, $P < .01$, which reflects the scoring scales' capacity to detect a clinically significant change after surgical intervention. Loss to follow-up ranged between 22% and 35% over the follow-up period. Therefore, sensitivity analyses were performed to assess the influence of nonresponders compared with a compliers-only analysis with a paired 2-sided Student's *t* test (Table 6).

Discussion

Scoring scale clinical instruments have traditionally been developed through a loose process that includes data thought to be clinically important to the investigator. However, this approach results in the genesis of scoring scales that generally provide insight into patient response to treatment but do not allow for comparison of results because they lack standard measurement and reporting techniques. The addition of objective data such as radiographic measurements with high interobserver reliability as well as reliable and valid joint range-of-motion techniques reduce investigator error and improve the usefulness of the scoring scale. In contrast to the above, the optimal scoring scale clinical instrument begins with a consensus panel that determines the critical elements to be assessed (the modified Delphi process). Once the areas of interest are determined, questions are crafted that will help to ascertain the relative success of the procedure or treatment to be evaluated. Draft scoring scales are then assembled and tested. Through constant refinement, the scale is eventually formed and must then be validated. The validation process is a formal, statistical process, and several specific criteria must be obtained. The optimal scoring scale clinical instrument will produce a quantity that is reliably reproducible and closely correlates with the patients' symptoms. This requires that the scoring scale has undergone a validation process that includes the following criteria:

- Validity based on content validity (face value), construct validity (subjective versus objective correlation), and criterion validity (correlation with gold standard)
- Reliability as demonstrated by consistency in data collection (intra-rater test–retest)
- Sensitivity to change after the study treatment

We conducted the first reevaluation of the ACFAS Scoring Scales to address the limitations of previous publications (4), to report the detailed methods used in their development, and to confirm the validation of modules 1 and 2 by evaluating the previously conducted (1–3) but not reported assessments of validity, reliability, and sensitivity to change. This was an extensive process that involved collection of all correspondence and materials archived by the ACFAS on the development of the original scoring scales and detailed review of this material over several sessions. We completely reviewed the previous work to assess the conclusions reported in the original publications and addressed the issues raised by others as mentioned previously. We reviewed the methods by which the original scoring scales were developed and the approach taken for instrument validation. Where possible, we determined standardization of the specific radiographic and objective functional items used in modules 1 and 2 to maximize reproducibility. To achieve this, we performed a detailed electronic search of the literature to determine if the techniques described possess a higher level of inter-rater and intra-rater reliability than those originally selected.

In the original publications of the ACFAS Scoring Scales (1–3), the validation process, data, and methods used were not reported. The subjective and objective sections were lacking in regard to definition, measurement criteria, and derivation of information. In addition, the

Table 2
Assessment of construct validity, modules 1 and 2: Patient reported outcomes by item

Question/Item				
Module 1: Subjective, n = 91	Preoperative	Postoperative	Expected Direction	Observed Direction
Pain				
None	7%	28%	↑	↑
Slight	13%	47%	↑	↑
Moderate	46%	19%	↓	↓
Significant	30%	6%	↓	↓
Severe	4%	0	↓	↓
Appearance				
Like very much	3%	30%	↑	↑
Mostly like	7%	32%	↑	↑
Neutral	25%	26%	↓	↑
Mostly dislike	37%	9%	↓	↓
Dislike	28%	2%	↓	↓
Function/Shoes				
Any shoe all the time	3%	23%	↑	↑
Any shoe most of the time	29%	40%	↑	↑
Only walking shoes	67%	36%	↓	↓
Only custom shoes	1%	2%	↓	↑
Module 1: Objective, n = 89	Preoperative	Postoperative	Expected Direction	Observed Direction
HA angle				
0°–20°	32%	82%	↑	↑
21°–30°	38%	7%	↓	↓
≥ 31°	2%	3%	↓	↑
–1°-3°	25%	7%	↓	↓
> –3°	3%	1%	↓	↓
IM angle				
0°–10°	28%	85%	↑	↑
11°–19°	66%	14%	↓	↓
≥ 20°	5%	0	↓	↓
< 0°	1%	1%	↓	—
First MT declination				
16°–24°	63%	68%	↑	↑
25°–29°	10%	24%	↑	↑
≥ 29°	5%	8%	↓	↓
10°–15°	23%	0	↓	↓
< 10°	N/A	N/A	N/A	N/A
Hallux purchase				
Not movable	39%	39%	↑	—
Resistant	39%	46%	↑	↑
Easy	22%	15%	↓	↓
ROM first MPJ DF				
≥ 60°	40%	40%	↑	—
45°–59°	22%	38%	↑	↑
36°–45°	20%	13%	↓	↓
< 36°	18%	10%	↓	↓
ROM first MPJ PF				
≥ 0°	91%	76%	↑	↓
< 0°	9%	24%	↓	↑
Hallux IPJ extension				
0°	98%	93%	↑	↓
< 0°	2%	7%	↓	↑
Limp: Yes	42%	19%	↓	↓
Module 2: Subjective, n = 84	Preoperative	Postoperative	Expected Direction	Observed Direction
Pain				
None	4%	26%	↑	↑
Slight	13%	38%	↑	↑
Moderate	42%	24%	↓	↓
Significant	33%	12%	↓	↓
Severe	8%	0	↓	↓
Appearance				
Like very much	12%	26%	↑	↑
Mostly like	18%	24%	↑	↑
Neutral	27%	32%	↓	↑
Mostly dislike	16%	15%	↓	↓
Dislike	27%	3%	↓	↓
Function/Shoes				
Any shoe all the time	6%	19%	↑	↑
Any shoe most of the time	23%	38%	↑	↑
Only walking shoes	68%	43%	↓	↓
Only custom shoes	4%	0	↓	↓

(*continued on next page*)

Table 2 (*continued*)

Question/Item				
Module 2: Objective, n = 82	Preoperative	Postoperative	Expected Direction	Observed Direction
4–5 IM angle				
0°–8°	84%	95%	↑	↑
≥ 9°	16%	5%	↓	↓
MT length	N/A	N/A	N/A	N/A
Transverse MPJ				
0°–5°	77%	91%	↑	↑
> 5°	23%	9%	↓	↓
Transverse IPJ				
0°–5°	77%	91%	↑	↑
> 5°	23%	9%	↓	↓
ROM MPJ DF				
≥ 65°	73%	82%	↑	↑
45°–64°	18%	16%	↓	↓
< 45°	9%	2%	↓	↓
ROM MPJ PF				
≥ 0°	88%	93%	↑	↑
< 0°	12%	7%	↓	↓
Digital purchase: Yes	67%	88%	↑	↑
Drawer sign				
Stable	73%	93%	↑	↑
Subluxable	16%	7%	↓	↓
Dislocated	11%	0	↓	↓
Limp: Yes	48%	0	↓	↓

Abbreviations: DF, dorsiflexion; HA, hallux abductus; IM, intermetatarsal; IPJ, interphalangeal joint; MPJ, metatarsophalangeal joint; MT, metatarsal; N/A, not applicable; PF, plantar flexion; ROM, range of motion.

rationale for the specific scaling (weighting) was not discussed. Despite these shortcomings, we have demonstrated that modules 1 and 2 of the ACFAS Scoring Scale (as originally published, i.e., without any additions or deletions) met the threshold for validation criteria as described in the last paragraph. Finally, through complete review of all pertinent information and detailed data analysis we have provided the evidence necessary to address the issues raised above.

There are myriad strengths of the ACFAS Scoring Scales for foot-related health measurement. The sample size for modules 1 and 2 was suitably large, and the design involved a multicenter patient enrollment process. The general ACFAS membership determined the desire to proceed with the development of the scoring scales with the specific intent of having this complete the rigors of validation. There was a strong assessment and evaluation of content (face) validity that we were able to achieve through independent review. A patient focus group was used and a large number of experts were consulted, as well as involved members of the original task force that used a modified Delphi approach. There was also an appropriate time span between the test–retest periods that allowed for assessment of reliability. Reliability, validity, and sensitivity to change were all tested, and small confidence intervals, indicating good precision of the reported results, were observed. Finally, to the authors' knowledge, modules 1 and 2 of the ACFAS Scoring Scales represent the first region-specific scoring scale that has been validated by quantifying validity, reliability, and sensitivity to change (6–16).

As with any meaningful publication, the strengths and message must be placed in the appropriate context by discussing identified limitations. Standardization of data collection methods was inconsistent for both subjective and objective components largely because of study intent to use the scale in multiple centers. Examples include patients completing the subjective scale in a waiting room versus in the presence of the surgeon. An objective component example would include angular measurement via digital rather than conventional radiographic films. Consistency in methodology is critical to internal validity and can decrease the accuracy of the data reported. In an effort to clarify and standardize data collection methods, Figure 6 has been constructed for reference when conducting the objective components of the first 2 modules. Although this is an important consideration, patients' results were compared with themselves at the different intervals and under similar conditions at those times. These differences are also reflective of the diversity of practice settings in which the scales were intended to be implemented into. Another limitation identified was in the definitions of measurements. In the original task force the numeric endpoints were developed via expert consultation, but a more robust approach is to use statistical methods to determine the distribution of measurements. Although not the ideal methodology, it is still considered an acceptable practice. Similarly, the weighting of questions was not determined by the optimal statistical methods but instead through a composite of consensus of expert consultation, patient focus groups, and others as noted above. As a result, question generation methodology was emphasized over question reduction. The scales therefore may include questions that are less useful and add to the overall burden of the data collection. Modules 3 and 4 have yet to undergo the necessary evaluations for validity, reliability, and sensitivity to change. Their use cannot be fully endorsed at this time; however, there are future plans to conduct these necessary assessments. Demographic data from the sample population were not archived by ACFAS and therefore could not be analyzed by the

Table 3
Construct validity, totals by parameter, modules 1 and 2

	Preoperative Total	Postoperative Total	Expected Direction	Observed Direction	*P* Value*
Module 1, subjective	21.8 ± 8.04	32.3 ± 11.04	↑	↑	.005
Module 1, objective	32.9 ± 10.03	38.8 ± 9.23	↑	↑	< .001
Module 2, subjective	20.8 ± 8.99	31.7 ± 9.32	↑	↑	< .001
Module 2, objective	30.4 ± 9.55	36.3 ± 4.47	↑	↑	< .001

* From 2-sided paired Student's *t* test. Patients with scores submitted for preoperative and postoperative testing.

Table 4
Reliability: Kappa coefficient for test–retest of subjective parameters by item, modules 1 and 2

	Preoperative Test–Retest	Response Rate	Postoperative Test–Retest	Response Rate
Module 1 pain	0.390[2]	74.7%	N/A	38.5%
Module 1 appearance	0.398[2]	74.7%	N/A	38.5%
Module 1 functional capacity	0.499[3]	74.7%	0.519[3]	38.5%
Module 2 pain	0.464[3]	76.1%	0.317[2]	38.0%
Module 2 appearance	0.512[3]	76.1%	0.509[3]	38.0%
Module 2 functional capacity	0.580[3]	76.1%	0.643[4]	38.0%

0 = poor agreement; 1 = slight agreement; 2 = fair agreement; 3 = moderate agreement; 4 = substantial agreement; 5 = almost perfect agreement (17).

Table 5
Reliability: Test–retest of subjective parameter totals, modules 1 and 2

	Test	95% CI	Retest*	95% CI
Preoperative mean module 1 score	21.8 ± 8.04	[20.13, 23.47]	23.2 ± 10.53	[20.67,25.73]
Postoperative mean module 1 score†	32.3 ± 11.04	[29.26, 35.34]	34.8.2 ± 9.9	[32.07, 37.53]
Preoperative mean module 2 score	20.8 ± 8.99	[18.85, 22.75]	22.4 ± 10.62	[19.77,25.03]
Postoperative mean module 2 score†	31.7 ± 9.32	[28.45, 34.95]	32.6 ± 11.45	[29.03, 36.17]

Abbreviation: CI, confidence interval.
* Retest administered 7 to 10 days after first test for operative period.
† Postoperative scores obtained 6 months after surgical intervention.

Table 6
Sensitivity for change: Outcomes of scales for preoperative, postoperative, and difference for modules 1 and 2

	Preoperative Total	Postoperative Total	Pre–Postoperative Difference*	Restricted Difference†	*P* Value*	95% CI*	*P* Value†	95% CI
Module 1, subjective	21.8 ± 8.04	32.3 ± 11.04	+4.7 ± 13.54	+13.5 ± 9.05	.005	[1.47, 7.93]	<.001	[9.99, 17.00]
Module 1, objective	32.9 ± 10.03	38.8 ± 9.23	+6.5 ± 11.28	N/A	< .001	[3.83, 9.17]	N/A	N/A
Module 2, subjective	20.8 ± 8.99	31.7 ± 9.32	+6.51 ± 11.26	+9.19 ± 10.17	< .001	[3.09, 9.93]	<.001	[5.46, 12.92]
Module 2, objective	30.4 ± 9.55	36.3 ± 4.47	+4.6 ± 8.26	N/A	< .001	[2.37, 6.83]	N/A	N/A

Abbreviations: CI, confidence interval; N/A, not applicable.
* Patients with scores submitted for preoperative and postoperative testing.
† Compliers only analysis with scores submitted for all 4 time periods.

authors. This makes generalizability difficult because the patient population cannot be strictly defined. Because multiple centers were selected, the samples are assumed to reflect an average foot and ankle surgeon's clinical environment. Cluster analysis would have been an ideal but onerous method for analyzing differences at each of the centers. Criterion validity was not assessed in this setting because a gold standard for comparison does not exist. Several other scales are routinely used in clinical research; however, comparison with them as a gold standard could not be assessed because of significant concerns related to their own validity, reliability, and sensitivity to change, or because they were not analogous. As per mandate, future reviews and updates will attempt to address these limitations to their fullest extent. These updates will include the addition and removal of questions, development of more global questions, and other issues or concerns as they are identified.

In conclusion, we have addressed the issues raised by reporting the detailed methods used in the development of the ACFAS Scoring Scales, as well as the steps used for prior validation of modules 1 and 2 that included assessments of validity, reliability, and sensitivity to change.

References

1. Zlotoff HJ, Christensen JC, Mendicino RW, Schuberth JM, Schwartz NH, Thomas JL, Weil LS Sr. ACFAS Universal Foot and Ankle Scoring System: first metatarsophalangeal joint and first ray (module 1). J Foot Ankle Surg 41(1):2–5, 2002.
2. Zlotoff HJ, Christensen JC, Mendicino RW, Schuberth JM, Schwartz NH, Thomas JL, Weil LS Sr. ACFAS Universal Foot and Ankle Scoring System: forefoot (module 2). J Foot Ankle Surg 41(2):109–111, 2002.
3. Thomas JL, Christensen JC, Mendicino RW, Schuberth JM, Weil LS Sr, Zlotoff HJ, Roukis TS, Vanore JV. ACFAS Scoring Scale user guide. J Foot Ankle Surg 44(5):316–335, 2005.
4. Lavery LA, Armstrong DG. Letter to the Editor: ACFAS Scoring Scale: ready, fire, aim? J Foot Ankle Surg 45(4):284–285, 2006.
5. Jones J, Hunter D. Consensus methods for medical and health services research. Br Med J 311(7001):376–380, 1995.
6. Schneider W, Knahr K. Scoring in forefoot surgery: a statistical evaluation of single variables and rating systems. Acta Orthop Scand 69(5):498–504, 1998.
7. Kitaoka HB, Patzer GL. Analysis of clinical rating scales for the foot and ankle. Foot Ankle Int 18(7):443–446, 1997.
8. SooHoo NF, Shuler M, Fleming LL. Evaluation of the validity of the AOFAS clinical rating systems by correlation to the SF-36. Foot Ankle Int 24(1):50–55, 2003.
9. Parker J, Nester CJ, Long AF, Barrie J. The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery. Foot Ankle Int 24(1):56–60, 2003.
10. 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 25(8):521–525, 2004.
11. 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 27(1):38–42, 2006.
12. Martin RL, Irrgang JJ, Lalonde KA, Conti S. Current concepts review: foot and ankle outcome instruments. Foot Ankle Int 27(5):383–390, 2006.
13. 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 27(11):930–934, 2006.
14. 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 27(12):1014–1019, 2006.
15. 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 46(2):65–74, 2007.
16. Van der Leeden M, Steultjens MPM, Terwee CB, Rosenbaum D, Turner D, Woodburn J, Dekker J. A systematic review of instruments measuring foot function, foot pain, and foot-related disability in patients with rheumatoid arthritis. Arthritis Rheum 59(9):1257–1269, 2008.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 33(1):159–174, 1977.