A new foot and ankle outcome score: Questionnaire based, subjective, Visual-Analogue-Scale, validated and computerized

Martinus Richter*, Stefan Zech, Jens Geerling, Michael Frink, Karsten Knobloch, Christian Krettek

Trauma Department, Hannover Medical School, Hannover, Germany

Received 3 October 2005; received in revised form 27 March 2006; accepted 4 April 2006

Abstract

Our purpose was to construct and validate a new score taking into consideration the flaws of existing scores.

Methods: A new score named Visual-Analogue-Scale Foot and Ankle (VAS FA) with the following features was constructed: questionnaire based on 20 subjective questions, Visual-Analogue-Scale (VAS) based rating, computerized evaluation. The score was validated in 121 subjects. For validation, SF-36® and Hannover Questionnaire (Q) were obtained and correlated with VAS FA.

Results: The correlation VAS FA versus SF-36® and Q (Pearson, all p-values < 0.001, r ≥ 0.5) was sufficient for the total score and all score categories (pain, function, other complaints).

The time needed for evaluating the scores was significantly lower for VAS FA than for SF-36® and Q (One-way-ANOVA, p < 0.001).

Conclusions: The introduced score is the first validated (on SF-36®), subjective, VAS based outcome score for foot and ankle. The VAS FA is computerized which enables faster evaluation than SF-36® or Q.

© 2006 European Foot and Ankle Society. Published by Elsevier Ltd. All rights reserved.

Keywords: Score; Visual-Analogue-Scale (VAS); Validation; Foot and ankle surgery

1. Introduction

Outcome assessment has become critical in evaluating the efficiency of both surgical procedures and medical treatments [1–3]. A wide variety of outcome measures have been proposed for use in conditions affecting the foot and ankle. A validated score for foot and ankle outcome is unavailable [1,4]. This deficiency was recently established at the American Orthopaedic Foot and Ankle Society’s (AOFAS) 2003 and 2005 Annual Summer Meetings [1–3]. Furthermore, the insufficiency of supposedly objective assessment was clearly stated [1]. Our purpose was to construct and validate a new score taking into consideration the flaws of existing scores.

There are two possibilities for validation:

(1) A score is validated for a specific population and language [5,6]. This requires an enormous case number, as for example more than 3000 subjects for the validation of the German SF-36® version [5]. Worldwide, this type of validation was achieved only for the SF-36® [5,6]. No foot and ankle score (including the AOFAS score) has ever been validated in that fashion.

(2) A score is validated by sufficient correlation with a validated score [7]. Since the SF-36® is the only validated score, this kind of validation is appropriate only by validation with that score [7]. Many scores for different types of diseases and populations have been validated following this principle [7–16]. Thus, no foot and ankle score (including the AOFAS score) has passed through this process as far as we know.
We choose the second type of validation process for this newly introduced score.

2. Methods

2.1. Validation process

Three scores (SF-36, Hannover Questionnaire (Q) and Visual-Analogue-Scale Foot and Ankle (VAS FA)) were obtained from a group of voluntary subjects. The results were analyzed and correlated to the scores. The time spent for evaluating the different scores was recorded and compared.

2.2. Subjects

One hundred and twenty-one subjects were included in the study. The number of subjects was determined by a statistician after a review of the study design and before starting the study. Fifty-eight of the subjects were female, 63 male, and the mean age was 31.4 [18–65] years. Exclusion criteria were as follows: age less than 18 or more than 65, positive medical history concerning the entire lower extremity, diabetes mellitus, drug abuse, psychiatric diseases and rheumatoid arthritis. The forms for the three scores (SF-36<sup>1</sup>, Q, VAS FA) were given to the subjects by one of the investigators (S.Z.) without any further explanation than provided by the score forms. The completed forms were

![Fig. 1. Visual-Analogue-Scale Foot and Ankle (VAS FA) form. English version shown for better understanding by the international reader. This version was translated from the German version that was used for this study. The form includes two pages, the instructions for completing the score on page 1 (a), and the 20 questions with the Visual-Analogue-Scale on page 2 (b).](image-url)
than evaluated by the same investigator who entered the score results into a personal computer for later statistical analysis.

2.3. Scores

2.3.1. SF-36®

The validated German version of the 36-item short-form (SF-36®) was used [5]. The SF-36® was developed to survey health status in the Medical Outcomes Study [5,6]. The SF-36® was designed for use in clinical practice and research, health policy evaluations and general population surveys. The SF-36® includes one multi-item scale that assesses eight health concepts:

1. limitations in physical activities because of health problems;
2. limitations in social activities because of physical or emotional problems;
3. limitations in usual role activities because of physical health problems;
4. bodily pain;
5. general mental health (psychological distress and well-being);
6. limitations in usual role activities because of emotional problems;
7. vitality (energy and fatigue);
8. general health perceptions. The SF-36® results were standardized to a possible maximum of 100 points to

<table>
<thead>
<tr>
<th>Strong</th>
<th>How much do foot problems affect your gait?</th>
<th>No changes, normal gait</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limping</td>
<td>How often do you have foot pain in physical rest?</td>
<td>Never, very rarely</td>
</tr>
<tr>
<td>Constantly, always</td>
<td>How intense is this foot pain in physical rest?</td>
<td>No pain</td>
</tr>
<tr>
<td>Extreme pain</td>
<td>How often do you have foot pain during physical activity?</td>
<td>Never, very rarely</td>
</tr>
<tr>
<td>Constantly, always</td>
<td>How strong is this foot pain during physical activity?</td>
<td>No pain</td>
</tr>
<tr>
<td>Extreme pain</td>
<td>Do you have the impression that one leg is weaker than the other?</td>
<td>Same strength as in the healthy leg</td>
</tr>
<tr>
<td>Widespread, painful callus</td>
<td>Do you have callous at the foot / feet?</td>
<td>No callus</td>
</tr>
<tr>
<td>My foot/ankle joint is constantly rigid</td>
<td>Do you have a limitation of ankle or foot range of motion?</td>
<td>No limitation of range of motion at any time</td>
</tr>
<tr>
<td>Climbing stairs impossible</td>
<td>Do you have problems when climbing stairs?</td>
<td>Climbing stairs without limitation possible</td>
</tr>
<tr>
<td>Occupation cannot be practiced any more</td>
<td>How much do foot problems affect your occupation?</td>
<td>No limitation</td>
</tr>
<tr>
<td>Driving a car not possible</td>
<td>How much do foot problems hinder you driving a car (operating clutch, accelerator, brake pedals)?</td>
<td>Driving a car without limitation possible</td>
</tr>
<tr>
<td>Only briefly, and with crutches/stick</td>
<td>How long can you stand without foot problems?</td>
<td>For hours, without limitation</td>
</tr>
<tr>
<td>Standing on one leg impossible</td>
<td>How long can you walk without foot problems?</td>
<td>No limitation</td>
</tr>
<tr>
<td>Impossible, or briefly with crutches/stick</td>
<td>Do foot problems stop you from running (e.g. jogging / on soft or uneven ground)?</td>
<td>Jogging for extended periods possible</td>
</tr>
<tr>
<td>Even short jogging is impossible</td>
<td>How much do foot problems affect your daily activities (e.g. getting dressed, eating, washing etc.)?</td>
<td>No limitation</td>
</tr>
<tr>
<td>Impossible on my own, need constant help</td>
<td>How much do foot problems restrict traveling (traveling with trains, busses, aircrafts etc.)?</td>
<td>No limitation</td>
</tr>
<tr>
<td>Traveling impossible</td>
<td>Do you have problems finding good footwear?</td>
<td>Can wear any type of shoe</td>
</tr>
<tr>
<td>Can only wear orthopaedic shoes</td>
<td>How much do foot problems restrict walking on uneven ground?</td>
<td>No limitations on uneven ground</td>
</tr>
<tr>
<td>On uneven ground walking is impossible</td>
<td>How much is your sensation in your foot / feet reduced?</td>
<td>Normal sensation</td>
</tr>
</tbody>
</table>

Fig. 1. (Continued).
allow a better comparison with the two other scores with a possible score maximum 100 points each.

2.3.2. Hannover Questionnaire

The Hannover Questionnaire rates patient’s complaints and the functional status based on a severity-symptom scale and functional status [17]. This score contains 20 questions for the patient with five possible answers to each. The questions require purely subjective answers.

2.3.3. Visual-Analogue-Scale Foot and Ankle (Fig. 1)

A new score named Visual-Analogue-Scale Foot and Ankle with the following features was constructed: a questionnaire based on 20 questions requiring purely subjective answers; three different question categories (pain, $n = 4$; function, $n = 11$; other complaints $n = 5$); Visual-Analogue-Scale (VAS) based rating; computerized evaluation. For each question a VAS-value from 0 to 100 points is possible. The total value for the entire score (all 20 questions answered) is therefore 0–2000 points. This total value is then divided by 20, resulting in a possible total score ranging from 0 to 100 points. To obtain the result from the single categories the total values from the category questions are divided by the number of questions (function 11; pain 4; other complaints 5). The different number of questions for the categories was determined to consider

Fig. 2. Template for evaluation of the Visual-Analogue-Scale Foot and Ankle (VAS FA) by hand (German version). The transparent template is placed on the score form and allows easy reading of the score values.
function, pain and other complaints with various import-
ances. Since more questions are included for function
\((n = 11)\) than for pain \((n = 4)\) and other complaints \((n = 5)\),
function is weighed higher for the final score than pain or
other complaints.

In case of missing answers the results of the entire score
or the categories can still be obtained by dividing the total
point value by the number of remaining questions. For
example, if one question is missing in each category (pain,
function, other complaints), the total value of the 17
remaining questions (ranging from 0 to 1700 points) is
divided by 17 to obtain the entire score. For the categories
(pain, function, other complaints), the total category values
of the remaining 3/10/4 questions is divided by 3/10/4 to
obtain the score category results.

The German score version was used for this study. Fig. 1
shows the English score version.

The score was evaluated visually and computerized.
Evaluation was performed with a transparent template that
was placed on the score form and allowed a reading of the
point values of the single questions (Fig. 2). The values were
then entered into a personal computer using a spreadsheet
based (Excel\textsuperscript{TM}, Microsoft Inc.) result-calculation-instru-
ment which enabled a calculation of the entire score result
and the category results. The computerized method included
a semi-automatic scanning of the score forms with a digital
tray (model Lightbox, Numonics Corporation, Montgo-
meryville, PA, USA) based computerized system (IBM
compatible personal computer; Windows 98\textsuperscript{®}, Microsoft
Inc.) (Fig. 3). Finally, the system calculated automatically
the entire score result and the category results.

The score forms, the template and the spreadsheet based
(Excel\textsuperscript{TM}, Microsoft Inc.) result-calculation-instrument for
evaluation have been posted on the author’s website
(German and English versions, DIN-A4 and letter format;

2.3.4. Score categories

The questions from the SF-36\textsuperscript{®} and Q were divided into
the categories pain, function and other complaints to allow a
correlation with these defined categories from the VAS FA.
Table 1 indicates the number of questions for each category
from all scores. For this part of the study, question 2 of the
SF-36\textsuperscript{®} (asking for a self assessment of changes in general
health compared to 1 year ago) was excluded from the
analysis because of a lack of comparable questions within
the other two scores. The total scores of the different
categories were standardized to a possible maximum of 100
points to allow better comparison between the scores.

2.4. Statistical analysis and hypothesis testing

Statistical analysis included One-way-ANOVA with Post
Hoc Scheffe\textsuperscript{®} test for time spent, and Pearson- or Spearman-
test for correlation of the score values. The total scores and
scores for each category (pain, function, and other
complaints) were correlated.

A null hypothesis at the \(p < 0.05\) level regarding time
spent of the score evaluation was formulated that there are
differences between the scores. A correlation was con-
sidered to be sufficient at a \(p < 0.05\) level and \(r > 0.5\) values.

A successful validation process of the VAS FA was
defined as sufficient correlation with the SF-36\textsuperscript{®} [7–16].

2.5. Ethical approval

The study was approved by the Ethical Commission of
the Hannover Medical School, Hannover, Germany.

<table>
<thead>
<tr>
<th>Score</th>
<th>Number of questions</th>
<th>Score values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Category pain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS FA</td>
<td>4</td>
<td>92.5</td>
</tr>
<tr>
<td>Q</td>
<td>4</td>
<td>95.7</td>
</tr>
<tr>
<td>SF-36\textsuperscript{®}</td>
<td>2</td>
<td>94.9</td>
</tr>
<tr>
<td>Category function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS FA</td>
<td>11</td>
<td>96.0</td>
</tr>
<tr>
<td>Q</td>
<td>9</td>
<td>98.4</td>
</tr>
<tr>
<td>SF-36\textsuperscript{®}</td>
<td>14</td>
<td>96.4</td>
</tr>
<tr>
<td>Category other complaints</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS FA</td>
<td>5</td>
<td>94.5</td>
</tr>
<tr>
<td>Q</td>
<td>7</td>
<td>96.2</td>
</tr>
<tr>
<td>SF-36\textsuperscript{®}</td>
<td>19</td>
<td>84.2</td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS FA</td>
<td>20</td>
<td>94.8</td>
</tr>
<tr>
<td>Q</td>
<td>20</td>
<td>96.7</td>
</tr>
<tr>
<td>SF-36\textsuperscript{®}</td>
<td>36</td>
<td>91.9</td>
</tr>
</tbody>
</table>

S.D., standard deviation.
Table 1
Results of the statistical correlation of SF-36®, Hannover Questionnaire (Q) and Visual-Analogue-Scale Foot and Ankle (VAS FA)

<table>
<thead>
<tr>
<th>Scores for correlation</th>
<th>Categories</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pain</td>
<td>Function</td>
</tr>
</tbody>
</table>
| VAS FA vs. Q           | $r = 0.5$, $p < 0.001$ | $r = 0.9$, $p < 0.001$ | $r = 0.5$, $p < 0.001$ | $r = 0.7$, $p < 0.001$
| VAS FA vs. SF-36®      | $r = 0.5$, $p < 0.001$ | $r = 0.6$, $p < 0.001$ | $r = 0.5$, $p < 0.001$ | $r = 0.6$, $p < 0.001$
| Q vs. SF-36®           | $r = 0.5$, $p < 0.001$ | $r = 0.7$, $p < 0.001$ | $r = 0.5$, $p < 0.001$ | $r = 0.7$, $p < 0.001$

Informed consent was obtained from all subjects included in the study.

3. Results

3.1. Score results and correlation

All questions were answered by all subjects.

Table 1 indicates the score results. Table 2 shows the results of the statistical correlation of the three scores. All correlations were significant regarding the defined $p < 0.05$ level. Sufficient correlation ($r > 0.5$) was found in all score categories and total scores.

The SF-36® scores did not significantly differ from the actual normative data (age and gender related) of the average German population (data not shown) [18].

3.2. Time spent

The time needed evaluating the scores was significantly lower for the VAS FA (evaluated visually or computerized) than for the SF-36® and Q (SF-36®, 244.5 (180–500) s; Q, 183.2 (145–420) s; VAS FA by hand 118.6 (90–180) s; VAS FA computerized 29.5 (22–60) s, Oneway-ANOVA, $p < 0.001$; Post Hoc Scheffé, VAS FA visually or computerized versus SF-36®, $p < 0.001/0.001$; VAS FA visually or computerized versus Q, $p < 0.001/0.001$). The null hypothesis regarding time spent was not rejected.

4. Discussion

The gold-standard score for foot and ankle is regarded as the AOFAS score [4,19–21]. This score is widely used as shown by the high number of hits when entering the term “aofas score” into the PubMed®-search engine in the World-Wide-Web (National Library of Medicine). However, this score is problematic due to significant flaws as follows. The score is not validated [1,4], it cannot be obtained if answers are missing, and contains problematic pseudo-objective assessment. To assess a walking distance in blocks, to specify joint stability as “stable” or “definitely instable”, to define gait abnormality as “none, slight”, “obvious” or “marked”, or to classify alignment as “good, plantigrade, well-aligned”, “fair, plantigrade, some degree of malalignment” or “poor, non-plantigrade, severe malalignment” are some examples for non- or pseudo-objective assessment of the AOFAS-score.

The Ankle Osteoarthritis Score (AOS) is an instrument that was developed for the assessment of pain in patients with ankle osteoarthritis [3,22]. The score demonstrated a high “vulnerability” regarding other musculoskeletal problems than at the ankle resulting in a questionable specifity for foot and ankle disorders [3].

The Foot Function Index (FFI) was lately correlated with the SF-36 for validation [2,23,24]. The correlation coefficients ranged from 0.10 to $–0.69$ for the different domains disability ($0.23$ to $–0.69$), activity limitation ($–0.26$ to $–0.64$) and pain ($–0.10$ to $–0.61$) [2]. Despite the conclusions of SooHoo et al. that these levels of correlation support the FFI as a valid measure, it is questionable that a correlation coefficient smaller than 0.5 (or greater than $–0.5$ for negative correlation) represents a sufficient correlation allowing a successful validation [2,25].

4.1. Construction of a new score

Based on these considerations we decided to construct and validate a new score. Only subjective assessment by the subjects was included to avoid pseudo-objective assessment. Consequently, a questionnaire based design with questions asking for a purely subjective answer by the subjects was created. The three different question categories pain, function, and other complaints comparable to the AOFAS score were found to be favourable for the new score [20]. These categories were weighed differently, because function was considered to be more important for the clinical outcome than pain alone or other complaints. More questions were included for function ($n = 11$) than for pain ($n = 4$) and other complaints ($n = 5$). Therefore, function is weighed higher for the final score than pain or other complaints. All questions, all categories and the entire score were designed to have the same range of possible points (0–100) for easy and distinct evaluation. Moreover, a high “stability” of the score against missing answers was expected. The current design allows us to obtain the entire score result and the result of the different score categories if answers are missing as described above.

Only Visual-Analogue-Scale based data acquisition was found to be appropriate. VAS has demonstrated to be the
most adequate technique for data acquisition regarding objectivity and reliability in numerous studies [26–32]. This technique is superior to category based evaluation as, for example in the AOFAS score or Q.

The evaluation process should be fast and simple. Therefore a simple method for evaluation with a template and a spreadsheet based result-calculation-instrument, and a computerized evaluation was requested.

Finally, the score forms and utilities for evaluation were planned to be available everywhere and free of charge, and were therefore posted on the author’s website (http://www.foot-trauma.org).

4.2. Scores for validation

For the validation of the new score, the short form 36 score (SF-36®) and the Hannover Questionnaire were chosen [5,6,17].

The SF-36® is a widely used and validated instrument that was constructed to satisfy psychometric standards necessary for group comparisons involving generic health concepts, that is, concepts that are not specific to any age, disease, or treatment group [5,6]. As a result, it may not be responsive to changes in a single organ system or a single body part as for example foot and ankle. The ability to measure responsiveness is a major need in functional measurement in chronic (non-fatal) disease [33,34]. Many countries have published so-called normative data from large sample sizes, which makes it possible to compare data from patients with different orthopedic or other conditions with data from the normal population of the country [21]. In summary, the SF-36® was used because it is the only widely validated score for outcome assessment.

The Hannover Questionnaire was chosen because it is an accepted questionnaire based subjective foot and ankle outcome score [17,35]. However, this score is not validated from a scientific point of view. The score does not use VAS-based data acquisition.

The AOFAS-score was not considered as an instrument for validation since it is not validated itself, does not use a VAS-based data acquisition, and contains problematic pseudo-objective assessment as described above [20].

4.3. Findings

We found a sufficient correlation (p < 0.05 and r > 0.5) between all three scores including the entire scores and the score categories. This means that all three scores assess the subjects in the same way. The sufficient correlation of the VAS FA with the validated SF-36® is the most important finding of the entire study since this specific finding completes the validation of the AOFAS successfully.

Regarding the time spent for the score evaluation, the VAS FA demonstrated faster evaluation than SF-36® and Q. In particular, the computerized evaluation of the VAS FA minimizes the time spent to approximately 30 s.

4.4. Methodological weaknesses

Subjects without known abnormalities were included in the validation process. This calls into question whether the introduced score is also valid for subjects with known foot and ankle abnormalities. However, the original validation process for the SF-36® also included “healthy” subjects [5,6]. The SF-36® scores of our group did not significantly differ from the actual normative data of the “healthy” German population [18]. Accordingly, an inclusion of “healthy” subjects for the validation of a foot and ankle score seems to be adequate. Furthermore, the results of all three scores for the entire scores and the different categories averaged approximately 95 of 100 possible points, reflecting the existence of foot complaints in this “healthy” group. Therefore, all three scores may be able to detect significant complaints as well.

The SF-36 form is not a disease-specific instrument and is not capable of detecting specific symptoms and limitations such as walking distances or restricted range of motion [21]. To review the sequelae of foot and ankle disorders, disease-specific or region-specific instruments should also be applied. Therefore, the Hannover Questionnaire as a non-validated but foot specific instrument was also included in the study. Finally, the VAS FA was validated in the German version and not in an – worldwide more useful – English version. Based on the experience with the SF-36®, the influence of the language on the validation process is low, and translated scores may be considered as validated as well without a new validation in each language [5,6].

The appropriate emphasis in function and pain is debatable. Consequently, we did not know how many of the questions should address pain, function of other complaints. We arbitrarily choose the number of questions for the different categories (function, 11; pain, 4; other complaints, 5) based on our experience with a VAS score that was earlier developed and validated for the spine [36]. Since all different categories successfully completed the validation process with the corresponding SF-36® categories, we believe the chosen “emphasis” in pain, function and other complaints is adequate.

We did not compare the VAS FA with the AOFAS score in the validation study but we currently use both scores in different follow-up studies. These studies deal with different foot and ankle topics such as osteochondral defects of the talus, arthrodeses of the ankle, hindfoot and midfoot, midfoot fractures and fracture-dislocations, Achilles tendon rupture and others. In these studies we used the AOFAS score for comparison with other studies. We consistently observe lower VAS FA scores than AOFAS scores (data not shown). Since the VAS FA is validated and the AOFAS score not, we consider the VAS FA score results as a more realistic outcome parameter.

In conclusion, the introduced score is the first validated (on SF-36®), subjective, VAS based outcome score for foot and ankle as far as we know. This score has important advantages in comparison with other scores. It does not include
problematic pseudo-objective assessment and/or problematic data acquisition (not VAS based). The introduced score is computerized which enables faster evaluation than SF-36\(^\text{[18]}\) or Q. The score form and the utilities for evaluation are available free of charge on the World Wide Web.

Acknowledgements

The authors thank Ludwig Hoy, Ph.D. (Institute for Biometry, Hannover Medical School, Hannover, Germany) for his help and support in carrying out the extensive statistical analysis and for his unbiased prior evaluation, Hans Werner Kuensebeck, M.D. (Department of Psychosomatics and Psychotherapy, Hannover Medical School, Hannover, Germany) for his assistance during the SF-36\(^\text{[18]}\) evaluation, and Justin Bender, Ph.D. for his assistance in translating the Visual-Analogue-Scale Foot and Ankle (VAS FA) form into the English version shown in Fig. 1, and for his review of the manuscript and the extensive language editing.

Appendix A. Links for download

Instruction

<table>
<thead>
<tr>
<th>Instruction</th>
<th>German version</th>
<th>English version</th>
</tr>
</thead>
</table>

References


