Systematic review of the psychometric properties of patient-reported outcome measures for rheumatoid arthritis in the foot and ankle

Ortega-Avila, AB, Ramos-Petersen, L, Cervera-Garvi, P, Nester, CJ, Morales-Asencio, J and Gijon-Nogueron, G

http://dx.doi.org/10.1177/0269215519862328

<table>
<thead>
<tr>
<th>Title</th>
<th>Systematic review of the psychometric properties of patient-reported outcome measures for rheumatoid arthritis in the foot and ankle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Ortega-Avila, AB, Ramos-Petersen, L, Cervera-Garvi, P, Nester, CJ, Morales-Asencio, J and Gijon-Nogueron, G</td>
</tr>
<tr>
<td>Type</td>
<td>Article</td>
</tr>
<tr>
<td>URL</td>
<td>This version is available at: <a href="http://usir.salford.ac.uk/id/eprint/51601/">http://usir.salford.ac.uk/id/eprint/51601/</a></td>
</tr>
<tr>
<td>Published Date</td>
<td>2019</td>
</tr>
</tbody>
</table>

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.
Systematic review of the psychometric properties of patient-reported outcome measures for rheumatoid arthritis in the foot and ankle

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Clinical Rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>CRE-2018-7845.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>18-Jun-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Ortega-Avila, Ana Belen; University of Malaga, Nursing and Podiatry Ramos-Petersen, Laura; University of Malaga, Nursing and Podiatry Cervera-Garvi, Pablo; University of Malaga, Nursing and Podiatry Nester, Christopher; University of Salford, Centre for Rehabilitation and Human Performance Research MORALES-ASENCIO, JOSE; University of Malaga, NURSING Gijon-Nogueron, Gabriel; University of Malaga, Nursing and Podiatry</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Rheumatoid arthritis, Foot and ankle, Psychometry, systematic review</td>
</tr>
</tbody>
</table>
Abstract

Objective. To identify self-reported outcome measures specific to the foot and ankle in patients with rheumatoid arthritis and to investigate the methodological quality and psychometric properties of these measures. Methods. A systematic review focusing on patients with rheumatoid arthritis. Setting: The search was conducted in the PubMed, SCOPUS, CINAHL, PEDro and Google Scholar databases, based on the following inclusion criteria: population (with rheumatoid arthritis) >18 years; psychometric or clinimetric validation studies of patient-reported outcomes specific to the foot and ankle, in different languages, with no time limit. Two of the present authors independently assessed the quality of the studies located and extracted the relevant data. Terwee’s criteria and the COSMIN checklist were employed to ensure adequate methodological quality. Results: Of the initial 431 studies considered, 14 met the inclusion criteria, representing 7,793 patients (56.8 years). These instruments were grouped into 3 dimensions (pain; perceived health status and quality of life and disability). The time to complete any of the PROMs varies around fifteen minutes. PROMs criterias with the worst scores by COSMIN, 92.85% and 85.71% were criterion validity, measurement error, internal consistency and responsiveness. 28.57% of PROMs were compared the measurement properties. Conclusion: the Self-Reported Foot and Ankle Score achieved the highest number of positive criteria (according to Terwee and COSMIN), and is currently the most appropriate for patients with Rheumatoid Arthritis

Keywords: Rheumatoid arthritis, Foot, Ankle, Psychometrics, Methodological quality, Patient-reported outcome measures, Measure.
Systematic review of the psychometric properties of patient-reported outcome measures for rheumatoid arthritis in the foot and ankle

Ana Belen Ortega-Avila1,2 PhD, Laura Ramos-Petersen1,2; Pablo Cervera-Garvi2 PhD; Christopher J Nester3 PhD; José Miguel Morales-Asencio2,4 PhD; Gabriel Gijon-Nogueron2,4 PhD

1 These authors contributed equally to the study/paper

2. Department of Nursing and Podiatry. Faculty of Health Sciences. University of Malaga, Spain

3. School of Health & Society, University of Salford, United Kingdom

4. Instituto de Investigación Biomédica de Málaga (IBIMA)

Corresponding author: anaortavi@uma.es

Department of Nursing and Podiatry. Faculty of Health Sciences. Arquitecto Francisco Penalosa 3. Ampliación de Campus de Teatinos, 29071 Málaga. Spain

Word count: 2354 words without references

Abstract: 222 words

Tables: 4

Figures: 2

Funding: This study did not receive any funding.

Conflict of interest: All the authors declare that they have no conflict of interest derived from the outcomes of this study.

Review registration number: PROSPERO (CRD42018090594).
Abstract

Objective. To identify self-reported outcome measures specific to the foot and ankle in patients with rheumatoid arthritis and to investigate the methodological quality and psychometric properties of these measures. Methods. A systematic review focusing on patients with rheumatoid arthritis. Setting: The search was conducted in the PubMed, SCOPUS, CINAHL, PEDro and Google Scholar databases, based on the following inclusion criteria: population (with rheumatoid arthritis) >18 years; psychometric or clinimetric validation studies of patient-reported outcomes specific to the foot and ankle, in different languages, with no time limit. Two of the present authors independently assessed the quality of the studies located and extracted the relevant data. Terwee’s criteria and the COSMIN checklist were employed to ensure adequate methodological quality. Results: Of the initial 431 studies considered, 14 met the inclusion criteria, representing 7,793 patients (56.8 years). These instruments were grouped into 3 dimensions (pain; perceived health status and quality of life and disability). The time to complete any of the PROMs varies around fifteen minutes. PROMs criterias with the worst scores by COSMIN, 92.85% and 85.71% were criterion validity, measurement error, internal consistency and responsiveness. 28.57% of PROMs were compared the measurement properties. Conclusion: the Self-Reported Foot and Ankle Score achieved the highest number of positive criteria (according to Terwee and COSMIN), and is currently the most appropriate for patients with Rheumatoid Arthritis

Keywords: Rheumatoid arthritis, Foot, Ankle, Psychometrics, Methodological quality, Patient-reported outcome measures, Measure.
Introduction

In patients with rheumatoid arthritis, foot pain, joint stiffness, deformity and loss of foot function are the major determinants of problems in foot-health-related quality of life (1-3). The consequences of foot problems in rheumatoid arthritis can be measured in a variety of ways, including physical activity (2), clinical status (3) and patient-reported outcome measures (4). The latter have the specific advantage of being meaningful to the individual patient, reflecting the issues that affect their health and lives. Existing patient-reported outcome measures differ in the foot-health concepts measured, but generally include pain (7-10), disability (8,10), function (5), activity limitation (7), footwear and general foot health (5).

In clinical practice, patient-reported outcome measures support physicians and patients, enabling them to co-create personalised care plans, taking into account patients’ preferences and values. For this purpose, robust instruments with good psychometric properties are necessary. Whilst many instruments for the foot and ankle are available (4), few are specific to rheumatoid arthritis (8,9), and their validation remains unclear. Further evidence is needed to determine how best to summarise and interpret the research data obtained and to determine the conditions that must be met in order to make well-founded recommendations. Furthermore, the evidence derived from research may be specific to the characteristics of the patients involved and rigorous methods are needed to overcome the potential bias associated with the study of human subjects.

The main aims of this review were to identify patient-reported outcome measures specific to the effects of rheumatoid arthritis in the foot and ankle, and to evaluate the methodological quality and psychometric properties of these instruments.
Material and Methods

This systematic review was carried out to assess patient-reported outcome measures used for patients with foot and ankle pathologies associated with rheumatoid arthritis. The review protocol was registered at the International Prospective Register of Systematic Reviews (PROSPERO: CRD 42018090594) prior to the identification of articles and data extraction.

Search strategy

The following databases were searched: PubMed, Scopus, CINAHL, PEDro and Google Scholar from inception until February 2018. All databases were searched again at the first of June 2019. In pubmed, the search was conducted in accordance with the strategy described by Terwee et al. (10) to detect the corresponding psychometric properties: construct search (patient-reported outcomes specific to the foot and ankle); population search (rheumatoid arthritis); instrument search (questionnaires, scales, instrument); measurement properties (filters). (Appendix 1).

The criteria applied for inclusion in the analysis were:

- Participants: patients with rheumatoid arthritis, aged over 18 years. The studies should be specifically focused on the foot and ankle;
- Studies: psychometric validation studies of patient-reported outcomes, published in English or Spanish;
- Outcomes: psychometric or clinimetric properties based on criteria according to Terwee (content validity; internal consistency; criterion validity; construct validity; reproducibility (agreement and reliability); responsiveness; floor/ceiling effect; interpretability) or COSMIN (structural validity; internal consistency; reliability; measurement error; hypothesis testing for construct validity; cross cultural validity/measurement invariance; criterion validity and responsiveness).
The exclusion criteria were:

- Studies: those based on questionnaires of orthopaedic injuries

**Quality appraisal**

The updated COSMIN checklist (Figure 1) was used to evaluate the methodological quality of studies investigating the measurement properties of a patient-reported outcome measure (11). This standard can be used either to assess the methodological quality of a study (12) or to compare the properties of various measurement instruments in a systematic review (13). The measurement properties considered are divided into three domains: reliability, validity and responsiveness. Each property contains various items, evaluated on a 4-point Likert scale as poor, fair, good or excellent. The “worst score counts” approach was applied to derive a final rating for each patient-reported outcome measure considered (13).

With respect to the psychometric properties proposed by Terwee (14), each issue was rated as positive “+” (adequate description or value or measure or argument related to the psychometric property), negative “-” (inadequate or values below the accepted standards for the psychometric property), indeterminate “?” (doubtful methods or measures or design) or absent “0” (no information available about the psychometric property), except for responsiveness, which was rated only as present/absent.

**Study selection**

Two blinded reviewers (LRP and PCG) evaluated the search results. The reference lists were reviewed independently to observe fulfilment or otherwise of the inclusion criteria. Disagreements were resolved by discussion between the two evaluators, or if consensus was not possible, further opinion was sought (ABOA, GGN, CN and JMMA).
Data extraction

Titles and abstracts were then reviewed independently by two reviewers (PCG and LRP) and relevant articles were then obtained in full text. The same reviewers undertook the second stage of screening by reading the full text of selected articles. The following data were extracted from each study, using a standardised template: full title, country, year of publication, dimensions and number of items, population used for the validation process, psychometric properties (Terwee’s criteria with a positive rating), cross-cultural adaptation into the language of each questionnaire included, and methodological quality (according to COSMIN). In studies lacking any of these elements, the authors were contacted to obtain the necessary data. The studies were first grouped into broad themes (according to the items), and then narrowed down into three main categories: pain, perceived health status/quality of life and disability.

No meta-analysis was carried out due to the heterogeneity of the dimensions and outcomes included in these studies.

Results

An initial 431 studies were identified, but 63 were duplicated among the different databases. The remaining 368 were screened against our inclusion/exclusion criteria, using the titles, abstracts and key words. Fifty seven studies met the inclusion criteria. After quality appraisal, a further 43 were excluded, and so 14 studies remained in the final analysis. Figure 2 shows the PRISMA flow diagram for the studies included in the review (15).

Population. A total of 7,793 participants were included in the 14 studies (61.4% female; 38.6% male, with a mean age of 56.8 years). The classification obtained for each measurement instrument is detailed in Table 1.
The Dimensions included in the different instruments were grouped into three areas (Table 2):

- pain (in the foot or ankle);
- perceived health status and quality of life (overall, lower limb-related or foot-related);
- disability (concerning activities of daily living, limitation of general function, limitation of sports/recreational function).

The range of dimensions were between two and seven. Four of the patient-reported outcome measures considered (the Ankle Osteoarthritis Scale, the Manchester Foot Pain and Disability Index, the Foot and Ankle Ability Measure and the Leeds Foot Impact Scale) had two dimensions, and one (the Podiatry Health Questionnaire) had seven dimensions.

Structure

The shortest patient-reported outcome measure (the Podiatry Health Questionnaire) had seven items, and the longest (the Leeds Foot Impact Scale) had 51.

Psychometric properties

The psychometric properties of each patient-reported outcome measure are summarised in Tables 1 and 2, following Terwee’s criteria. The Self-reported Foot and Ankle Score, included in the pain group, presented the best overall psychometric properties, with positive evidence for content validity (clear description of measurement aim, target population, item selection and reduction), internal consistency (Cronbach’s alpha 0.70-0.95), construct validity (evidence from factor analysis to confirm the study hypotheses), reproducibility/reliability (ICC>0.7), floor/ceiling effect (only described for the Self-reported Foot and Ankle Score (0%)). On the other hand, the evidence was indeterminate
for three criteria (reproducibility:agreement, responsiveness and interpretability) and negative for one (criterion validity).

In the perceived health status/quality of life group, there was positive evidence for the Foot Health Status Questionnaire on four criteria: content validity, internal consistency, construct validity and reproducibility:reliability.

In the disability group, there was positive evidence for the Rheumatoid and Arthritis Outcome Score on three criteria: content validity, internal consistency and reproducibility/reliability.

With respect to criterion validity; reproducibility:agreement, responsiveness and interpretability, positive ratings were obtained in very few cases; most of the patient-reported outcome measures considered obtained an indeterminate or absent rating.

**Cross-Cultural Adaptation**

Neither the Rowan Foot Pain Assessment Questionnaire nor the Salford Rheumatoid Arthritis Foot Evaluation considered the question of cross-cultural adaptation. The other patient-reported outcome measures had been translated or culturally adapted into diverse languages, including Arabic, Somali, Thai, Danish, Spanish, Hungarian, Polish and Greek. In this respect, the Foot and Ankle Ability Measure was the most widely adapted, being translated into eleven languages (French, Japanese, Persian, German, Italian, Turkish, Brazilian, Spanish, Chinese, Thai and Dutch).

**Methodological Quality**

The Self-reported Foot and Ankle Score and the Foot and Ankle Ability Measure were assessed by the COSMIN criteria for methodological quality (Table 3). The first of these patient-reported outcome measures had a positive rating for reliability, hypothesis testing for construct validity and responsiveness, a negative one for structural validity and criterion validity, and indeterminate ratings for internal consistency, measurement error
and cross-cultural validity/measurement invariance. The second had a positive rating for reliability, measurement error and responsiveness, a negative one for structural validity, hypothesis testing for construct validity and criterion validity, and indeterminate ratings for internal consistency and cross-cultural validity/measurement invariance. Overall, both presented poor methodological quality.

For the following properties, the other patient-reported outcome measures had few positive ratings, often presenting missing or unknown data: internal consistency (Cronbach’s alpha not determined or dimensionality unknown), measurement error (patient-reported outcome measures not defined by minimally-important change), hypothesis testing (hypothesis not defined or results conflicting with the hypothesis), cross-cultural/measurement invariance (no important differences found between group factor or differential item functioning), criterion validity or responsiveness (no hypothesis defined, results conflicting with the hypothesis or area under the curve <0.70)

- Methodological quality according to measurement properties

In addition to the above, we evaluated the methodological quality of the best-rated patient-reported outcome measures, using COSMIN boxes to classify their quality as poor, fair, good or excellent. These details are shown in Table 4. In this respect, only the Foot Health Status Questionnaire, the Foot and Ankle Ability Measure, Salford Rheumatoid Arthritis Foot Evaluation and the Self-reported Foot and Ankle Score achieved a positive score according to COSMIN. In the context of the low overall score, the Foot and Ankle Ability Measure was rated highest, with excellent ratings for content validity, structural validity and criterion validity. None of these patient-reported outcome measures were evaluated for cross-cultural validity as the inclusion criteria limited the studies considered to those focusing on rheumatoid arthritis.
Discussion

The objective of this systematic review was to identify patient-reported outcome measures concerning the effects of rheumatoid arthritis on the foot and ankle, and to evaluate the methodological quality and psychometric properties of these measures. The Self-reported Foot and Ankle Score presented the best overall psychometric properties and methodological quality. With respect to psychometric properties, the Self-reported Foot and Ankle Score (16) obtained the highest number of positive criteria, although it presented deficiencies in criterion validity, agreement, responsiveness and interpretability. This patient-reported outcome measure is relatively new and to date only one cross-cultural adaptation (into German) has been made.

The patient-reported outcome measures analysed in this review had 2-7 dimensions and were further categorised into three areas: pain, perceived health status and quality of life and disability, according to their main components. Similar categorisations have been performed by Van der Leeden et al. (4) and Oude Voshaar et al. (2), both of whom combined patient-reported outcome measures with scales and other instruments measuring foot function, pain or foot-related disability.

Most of the patient-reported outcome measures analysed have been culturally adapted for use in other languages. Such transcultural adaptations are important, enabling health professionals in different societies and countries to have the same perspective and to obtain comparable data for patients with rheumatoid arthritis. On the other hand, if it is to be valid, any such cross-cultural adaptation must be performed with scientific rigour.

Most of the patient-reported outcome measures considered presented deficiencies regarding construct validity, responsiveness, floor/ceiling effect and interpretability. It is important to highlight these shortcomings, as they may have significant consequences in clinical and research contexts. Construct validation is an on-going process of learning,
prediction and testing (17). If it is not performed appropriately, the resulting conclusions on assisting patients in the development of self-management skills will be unreliable and discounted.

Another important question is that of the floor/ceiling effect. This parameter helps identify any redundant items it may include. Obviously, if a patient-reported outcome measure did not provide information about what (change in) score would be clinically meaningful, it would have little practical or theoretical value.

The study presents certain limitations. Importantly, some instruments were excluded from our analysis, namely the Oxford Ankle Foot Questionnaire for Children (18) and the Juvenile Arthritis Foot Disability Index (19), due to our focus on patients aged over 18 years, therefore, our findings could only be related to adult RA population. Another limitation was the fact that some data were incomplete, despite our efforts to contact the original authors. Among its strengths, this study was based on a literature search of five medical databases, with a well-defined search strategy and no limitation on time. Moreover, all the studies included had been clinimetrically validated. The review we describe was based on a blinded quality appraisal following a well-established method, the COSMIN checklist.

The clinical implications of these results point out the gap regarding the dimension of self-care, prevention or treatment adherence specifically with respect to the foot and ankle. This issue is of major importance to patients with rheumatoid arthritis, as its impact on the foot and ankle often limits or prevents the activities of daily life. Instruments with these dimensions should be available for patients and clinicians.

On the other hand, the scarcity of responsiveness evaluation for most of the instruments implies a major shortfall for clinical practice. The criterion of responsiveness is of crucial importance, revealing the clinically important changes that must be observed and helping
clinicians and patients monitor the condition. Moreover, this issue may jeopardize the outcome evaluation in longitudinal research.

Future research should address the structure of the questionnaires considered; the number of items varied widely among the patient-reported outcome measures, and response options were also heterogeneous, with some offering a simple yes/no choice, while others measured outcomes on a Likert scale. In future research, it would be useful to examine whether the number of items and the response options provided correctly discriminate the interventions performed, the health status of the patients and the follow up procedures employed.
Clinical messages

1. On available evidence, the **Self-reported foot and ankle score** is currently the most appropriate patient-reported outcome measure available for patients with Rheumatoid Arthritis.

2. The most of patient-reported outcome measures have poor evidence of their psychometric properties and should be used with caution for patients with Rheumatoid Arthritis.

3. Robust methods should be designed and implemented to get higher-quality instruments for patients with Rheumatoid Arthritis.
References


trez&rendertype=abstract


an


69. Ling SKK, Chan V, Ho K, Ling F, Lui TH. Reliability and validity analysis of the open-source Chinese Foot and Ankle Outcome Score (FAOS). Foot. 2018;

Tables and Figure

Figure 1. Instruction for completing the COSMIN checklist

Figure 2. PRISMA. Flow diagram

Table 1. Instruments included in the study

Table 2. Assessment of the measurement properties of the questionnaires

Table 3. Detailed COSMIN ratings

Table 4. Methodological quality per PROM property (COSMIN)
<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Data of Psychometric properties</th>
<th>Dimensions and items</th>
<th>Population used for validation</th>
<th>Psychometric properties</th>
<th>Cross-cultural adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. Budiman-Mak et al. 1991</td>
<td>Internal consistency: Cronbach’s alpha 0.88-0.94 Test-retest reliability: (0.64-0.79) 2 factors: ICC (0.70-0.83) – ICC (0.63-0.71)</td>
<td>3 dimensions: pain, disability and activity restriction 23 items</td>
<td>87 patients with RA 77 male (89%) 10 female (11%) Mean age: 61 years (24-79)</td>
<td>Internal consistency: Cronbach’s alpha 0.96-0.73 (total: 0.95) Test-retest reliability: (0.87 – 0.69). ICC= 0.87 4 factors: foot pain (1-9) disability (10-18) activity limitation (19-21) social issues (22-23)</td>
<td>8 Brazilian/Portuguese(20), Polish (20), Korean(21), Italian(22), Taiwan Chinese(23), French(24), Spanish (25), German(26)</td>
</tr>
<tr>
<td>E. Budiman-Mak et al. 2006</td>
<td>Internal consistency: Cronbach’s alpha 0.88-0.94 Test-retest reliability: (0.64-0.79) 2 factors: ICC (0.70-0.83) – ICC (0.63-0.71)</td>
<td>18 items</td>
<td>562 patients 264 male (47%) 298 female (53%) Age 20-85 years</td>
<td>Test - retest analysis ICC of 0.97 (0.94-0.99)</td>
<td>1 French(29)</td>
</tr>
<tr>
<td>P. J. Bennett et al. 1998</td>
<td>Internal consistency: Cronbach’s alpha between 0.85 and 0.88</td>
<td>4 dimensions: foot pain, foot function, footwear, and general foot health</td>
<td>111 patients 25 male (22.5%)</td>
<td>Internal consistency: Cronbach’s alpha between 0.85 and 0.88</td>
<td>2 Spanish(30), Brazilian(31)</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Items</td>
<td>Mean Age</td>
<td>Gender</td>
<td>Construct Validity</td>
<td>Reliability</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------</td>
<td>----------</td>
<td>--------</td>
<td>--------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Foot Health Status Questionnaire</td>
<td>13</td>
<td>45</td>
<td>85 (77.5%)</td>
<td>4 factors</td>
<td>ICC 0.74-0.91</td>
</tr>
<tr>
<td>MFPDI(6) Manchester Foot Pain Disability Index</td>
<td>2 dimensions: foot pain and disability</td>
<td>1078</td>
<td>604 (56%)</td>
<td>6.42-34.9%</td>
<td>kappa 0.48, 0.50, 0.17</td>
</tr>
<tr>
<td>ROFPAQ(36) Rowan Foot Pain Assessment Questionnaire</td>
<td>3 dimensions: multi-dimensional pain (sensory-discriminative, motivational-affective and cognitive-evaluative)</td>
<td>17</td>
<td>5 (29%)</td>
<td>Cronbach’s α between 0.80 and 0.90</td>
<td>Spearman correlations with Headache scale from 0.15 to 0.48</td>
</tr>
<tr>
<td>Test Code</td>
<td>Test Name</td>
<td>Authors</td>
<td>Year</td>
<td>Description</td>
<td>Sample Size</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PHQ(37)</td>
<td>Podiatry Health Questionnaire</td>
<td>S. Macran et al.</td>
<td>2003</td>
<td>7 dimensions: walking, hygiene, nail care, foot pain, worry/concern, quality of life and PHQ&lt;sub&gt;vas&lt;/sub&gt;</td>
<td>2073 patients</td>
</tr>
<tr>
<td>RAOS(37)</td>
<td>Rheumatoid and Arthritis Outcome Score</td>
<td>A BI. Bremander et al.</td>
<td>2003</td>
<td>5 dimensions: Pain; other symptoms like stiffness, swelling, and range of motion; activities of Daily Living (ADL); sport and Recreational activities (Sport/Rec); and lower limb-related Quality of Life (QOL).</td>
<td>119 patients with inflammatory joint disease (51% RA)</td>
</tr>
<tr>
<td>FAM-AAOS(42)</td>
<td>Foot and Ankle Module of American Academy of Orthopaedic Surgeons</td>
<td>N A. Johanson et al.</td>
<td>2004</td>
<td>5 dimensions: function, pain, stiffness and swelling, giving way and shoe comfort</td>
<td>205 patients</td>
</tr>
</tbody>
</table>

1. Spanish(38)
2. Turkish(39)
3. French (40) Persian(41)
<table>
<thead>
<tr>
<th>Scale</th>
<th>Authors</th>
<th>Year</th>
<th>Items</th>
<th>Subscales/Dimensions</th>
<th>Patients</th>
<th>Gender</th>
<th>Mean Age</th>
<th>Construct Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAAM(44)</td>
<td>R.L. Martin et al.</td>
<td>2005</td>
<td>21</td>
<td>2 dimensions: activities of daily living (ADL) and sports</td>
<td>1027</td>
<td>391</td>
<td>42</td>
<td>Criterion validity: with SF-36 function subscale (r = 0.84, 0.78), physical component summary score (r = 0.78, 0.80), mental function subscale (r = 0.18, 0.11) and mental component summary score (r = 0.05, -0.02). Construct validity: one factor in Group 1 (80.46% of the variance and an eigenvalue of 16.90). Two factors in Group 2 (first factor 78.37% of the variance and an eigenvalue of 16.46; second factor 12.28% of the variance and an eigenvalue of 2.58). Agreement: minimal detectable change for the ADL subscale ±5.7. For the Sports subscale ±12.3 points. Minimal clinically important difference for ADL 8 and for Sports subscale 9 points. Test-retest reliability: 4 weeks apart. 0.89 and 0.87 for the ADL and Sports subscales, respectively.</td>
</tr>
<tr>
<td>BFS(56)</td>
<td>S. Barnett et al.</td>
<td>2005</td>
<td>400</td>
<td>5 dimensions: mobility, pain, footwear, foot health and disability,</td>
<td>1027</td>
<td>391</td>
<td>42</td>
<td>Internal consistency: Cronbach’s α= 0.90</td>
</tr>
</tbody>
</table>
| Bristol Foot Score | and perception of self as a result of foot problems | 3 male (30%)  
7 female (70%)  
Age 24 to 89 years  
Version 4 71  
23 male (32%)  
48 female (68%)  
Mean age 58 years (13-90) | 3 factors: feet pain (50%), footwear and general foot health (10%) and mobility (9%). |
|---|---|---|
| LFIS(58) Leeds Foot Impact Scale | P. Helliwell et al.  
2005 | 2 dimensions:  
impairment/shoe and activities/participation  
51 items | 192 patients with RA  
yielded 148  
34 male (23%)  
114 female (77%)  
Mean age 61.7 years (28-89) | Content validity: qualitative pilot study with 30 subjects  
Reliability: Impairment / shoes subscale ICC of 0.84 (95% CI 0.75–0.90); Activities / participation subscale ICC of 0.96 (95% CI 0.93–0.98).  
3  
Dutch(59)  
German  
Hungarian(60) |
| SAFE(61) Salford Rheumatoid Arthritis Foot Evaluation | S. Walmsley  
2012 | 3 dimensions:  
impairment, disability and foot wear  
19 items | 28 patients  
7 male (25%)  
21 female (75%)  
Mean age 58,5 | Content validity: qualitative study  
Criterion validity: MFPDI 0.83 and LFIS 0.79  
0 |
| FAOS(62) Foot and Ankle Outcome Score | Y M. Golightly et al. 2014 | 1670 patients  
541 male (32.4%)  
1129 female (67.6%)  
Mean age 69 years (50-95)  
Group 1 (pain) 1641  
Group 2 (ADL) 1609  
Group 3 (sport/recreation) 1454  
Group 4 (QOL) 1632  
Group 5 (other symptoms) 1670 | Internal consistency: group 1 Cronbach’s α= 0.95 – 0.97; group 2 Cronbach’s α= 0.97-0.98; group 3 Cronbach’s α= 0.94 – 0.96; group 4 Cronbach’s α= 0.89 – 0.92; group 5 Cronbach’s α= 0.72 – 0.82  
Reliability: ICC=0.63 – 0.81 | Persian(63) Korean(64) Dutch(65) German(66) Thai(67) Turkish(68) Chinese(69) |
<table>
<thead>
<tr>
<th>SEFAS (16)</th>
<th>M. Cöster et al. 2014</th>
<th>224 patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported Foot and Ankle Score</td>
<td>3 dimensions: pain, function, and limitation of function</td>
<td>12 items</td>
</tr>
<tr>
<td></td>
<td>224 patients</td>
<td>Group 1 (Foot disorders):</td>
</tr>
<tr>
<td></td>
<td></td>
<td>118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22 male (19%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>96 female (81%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean age 57 years (16–87)</td>
</tr>
<tr>
<td></td>
<td>Group 2 (midfoot, hindfoot or ankle disorders):</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td></td>
<td>47 male (44%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59 female (56%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean age 55 years (18–81)</td>
</tr>
<tr>
<td>Internal consistency: group 1 Cronbach’s α = 0.84; group 2 Cronbach’s α = 0.86,</td>
<td>Criterion validity: Spearman rho with FAOS, SF-36, EQ-5D (0.6 – 0.8)</td>
<td></td>
</tr>
<tr>
<td>Construct validity: 80% of predefined hypotheses confirmed</td>
<td>Reliability: group 1 ICC = 0.92; group 2 ICC = 0.93</td>
<td></td>
</tr>
<tr>
<td>Floor/ceiling effect: group 1 = 0%; group 2 = 0%</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

RA Rheumatoid Arthritis; N number of patients; ICC Intraclass correlation coefficient; ADL Activities of Daily Living; SF-36 Short Form-36 health survey; EQ-5D EuroQol-5D
Table 2. Assessment of the measurement properties of the questionnaires

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Content validity</th>
<th>Internal consistency</th>
<th>Criterion validity</th>
<th>Construct validity</th>
<th>Reproductibility Agreement</th>
<th>Reproductibility Reliability</th>
<th>Responsiveness</th>
<th>Floor/ceiling effect</th>
<th>Interpretability</th>
<th>Final assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOS</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>?</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MFPDI</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ROFP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEFAS</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>V</td>
</tr>
<tr>
<td>FHSQ</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>PHQ</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>BFS</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>FAOS</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Perceived Health Status and Quality of Life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAAM</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FF1</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>-</td>
<td>?</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>RAOS</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>?</td>
<td>V</td>
</tr>
<tr>
<td>RAOS</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>FAAM</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

http://mc.manuscriptcentral.com/clinrehab
<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAM AAOS</td>
<td>+ + - 0 0 - 0 0</td>
</tr>
<tr>
<td>LFIS</td>
<td>+ 0 ? 0 0 + 0 ?</td>
</tr>
<tr>
<td>SAFE</td>
<td>+ 0 + 0 ? + 0 ?</td>
</tr>
</tbody>
</table>

Group:
- Pain: AOS Ankle Osteoarthritis Scale; MFPDI Manchester Foot Pain Disability Index; ROFPAQ Rowan Foot Pain Assessment Questionnaire; SEFAS Self-reported Foot and Ankle Score
- Perceived Health Status and Quality of Life: FHSQ Foot Health Status Questionnaire; PHQ Podiatry Health Questionnaire; BFS Bristol Foot Score; FAOS Foot and Ankle Outcome Score
- Disability: FFI Foot Function Index; RAOS Rheumatoid and Arthritis Outcome Score; FAAM Foot and Ankle Ability Measure; FAM-AAOS Foot and Ankle Module of American Academy of Orthopaedic Surgeons; LFIS Leeds Foot Impact Scale; SAFE Salford Rheumatoid Arthritis Foot Evaluation

Rating: + Positive; ? Indeterminate; - Negative; 0 No information available.
Table 3. COSMIN ratings

<table>
<thead>
<tr>
<th></th>
<th>Structural Validity</th>
<th>Internal Consistency</th>
<th>Reliability</th>
<th>Measurement Error</th>
<th>Hypothesis testing for construct validity</th>
<th>Cross Cultural Validity/Measurement Invariance</th>
<th>Criterion Validity</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHSQ</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>FAAM</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>?</td>
</tr>
<tr>
<td>SAFE</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>SEFAS</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Rating: “+”: Positive; “?”: Indeterminate; “-”: Negative
FFI Foot Function Index; AOS Ankle Osteoarthritis Scale; FHSQ Foot Health Status Questionnaire; MFPDI Manchester Foot Pain Disability Index; ROFPAQ Rowan Foot Pain Assessment Questionnaire; PHQ Podiatry Health Questionnaire; RAOS Rheumatoid and Arthritis Outcome Score; FAM-AAOS Foot and Ankle Module of American Academy of Orthopaedic Surgeons; FAAM Foot and Ankle Ability Measures; BFS Bristol Foot Score; LFIS Leeds Foot Impact Scale; SAFE Salford Rheumatoid Arthritis Foot Evaluation; SEFAS Self-reported Foot and Ankle Score; FAOS Foot and Ankle Outcome Score
Table 4. Methodological quality per PROM property (COSMIN)*

<table>
<thead>
<tr>
<th>BOX A</th>
<th>BOX B</th>
<th>BOX C</th>
<th>BOX D</th>
<th>BOX E</th>
<th>BOX F</th>
<th>BOX G</th>
<th>BOX H</th>
<th>BOX I</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHSQ</td>
<td>Fair</td>
<td>Poor</td>
<td>Fair</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>-</td>
<td>Poor</td>
</tr>
<tr>
<td>FAAM</td>
<td>Poor</td>
<td>Poor</td>
<td>Good</td>
<td>Excellent</td>
<td>Excellent</td>
<td>Good</td>
<td>-</td>
<td>Excellent</td>
</tr>
<tr>
<td>SAFE</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Excellent</td>
<td>Poor</td>
<td>Poor</td>
<td>-</td>
<td>Poor</td>
</tr>
<tr>
<td>SEFAS</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Excellent</td>
<td>Poor</td>
<td>Fair</td>
<td>-</td>
<td>Poor</td>
</tr>
</tbody>
</table>

* COSMIN checklist can be used to assess the quality of a study on one measurement instrument or to compare the measurement properties of a number of measurement instruments in a systematic review.

FHSQ: Foot Health Status Questionnaire; FAAM Foot and Ankle Ability Measures; SAFE Salford Rheumatoid Arthritis Foot Evaluation; SEFAS Self-reported Foot and Ankle Score
**Figure 1.** Instructions for completing the COSMIN checklist

---

**INSTRUCTIONS FOR COMPLETING THE COSMIN CHECKLIST**

1. **Mark the properties that have been assessed in the article.**
   - A. Internal consistency
   - B. Reliability
   - C. Measurement error
   - D. Content validity
   - E. Construct validity
   - F. Structural validity
   - G. Hypothetico-deductive validity
   - H. Cross-cultural validity
   - I. Criterion validity
   - J. Responsiveness
   - K. Interpretability

2. **Are IRT methods used in the article?**
   - Yes
   - No

3. **Complete for each property you marked in step 1 the corresponding box A to J.**

4. **Complete for each property you marked in step 1 the Generalisability box.**

---

http://mc.manuscriptcentral.com/clinrehab
Figure 2. PRISMA Flow Diagram

Records identified by database search (n = 431)

Additional records identified from other sources (n=0)

Duplicates removed (n=63)

Records screened (title and abstract) (n = 368)

Full-text articles assessed for eligibility (n = 57)

Studies included in qualitative synthesis (n = 14)

Records excluded (Psychometric validation studies of PROMs not focused on foot and ankle) (n =311)

Records excluded for methodological reasons (n =43)
### Appendix 1. Searching Strategy (Pubmed)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rheumatoid Arthritis</td>
</tr>
<tr>
<td>2</td>
<td>Foot</td>
</tr>
<tr>
<td>3</td>
<td>Feet</td>
</tr>
<tr>
<td>4</td>
<td>Ankle</td>
</tr>
<tr>
<td>5</td>
<td>2 OR 3 OR 4</td>
</tr>
<tr>
<td>6</td>
<td>1 AND (2 OR 3 OR 4)</td>
</tr>
<tr>
<td>7</td>
<td>“Patient Reported Outcome Measures”</td>
</tr>
<tr>
<td>8</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>9</td>
<td>Instrument</td>
</tr>
<tr>
<td>10</td>
<td>Scale</td>
</tr>
<tr>
<td>11</td>
<td>Index</td>
</tr>
<tr>
<td>12</td>
<td>7 OR 8 OR 9 OR 10 OR 11</td>
</tr>
<tr>
<td>13</td>
<td>6 AND 12</td>
</tr>
<tr>
<td>14</td>
<td>“Pain”</td>
</tr>
<tr>
<td>15</td>
<td>Disab*</td>
</tr>
<tr>
<td>16</td>
<td>Funct*</td>
</tr>
<tr>
<td>17</td>
<td>14 OR 15 OR 16</td>
</tr>
<tr>
<td>18</td>
<td>13 AND 17</td>
</tr>
</tbody>
</table>

#1 Rheumatoid Arthritis [tiab] 100484
#2 Foot [tiab] 90109
#3 Feet [tiab] 27300
#4 Ankle [tiab] 54011
#5 ((foot [tiab]) OR feet [tiab]) OR ankle [tiab]) 147500
#6 (rheumatoid arthritis[tiab]) AND (((foot [tiab]) OR feet [tiab]) OR ankle [tiab]) 2387
#7 "Patient Reported Outcome Measures"[Mesh] 3291
#8 questionnaire[tiab] 373536
#9 instrument [tiab] 110243

#10 scale [tiab] 652290

#11 index [tiab] 713203

#12 ((("Patient Reported Outcome Measures"[Mesh]) OR questionnaire[tiab]) OR instrument [tiab]) OR scale [tiab]) OR index [tiab] 1663798

#13 (((rheumatoid arthritis[tiab]) AND (((foot [tiab]) OR feet [tiab]) OR ankle [tiab]))) AND ((("Patient Reported Outcome Measures"[Mesh]) OR questionnaire[tiab]) OR instrument [tiab]) OR scale [tiab]) OR index [tiab]) 524

#14 "Pain"[Mesh] 375794

#15 disab* 328598

#16 funct* 370020

#17 (pain[Mesh]) OR funct* OR disab* 4265399

#18 (((rheumatoid arthritis[tiab]) AND (((foot [tiab]) OR feet [tiab]) OR ankle [tiab]))) AND ((("Patient Reported Outcome Measures"[Mesh]) OR questionnaire[tiab]) OR instrument [tiab]) OR scale [tiab]) OR index [tiab]) AND ((pain[Mesh]) OR funct*) OR disab*) 262