Development of Patient-Reported Outcome Scale for Patients with Diabetic Foot and Its Reliability and Validity Test

Xuanyu Wang*, Xiaojie Hu*, Huafa Que

Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, 200032, People's Republic of China

*These authors contributed equally to this work

Correspondence: Huafa Que, Email huafaque@126.com

Objective: To construct a self-reported outcome scale for diabetic foot patients, and to test its reliability and validity.

Methods: Through literature reading and interviews with 30 patients, a pool of scale items was formed. The items were classified and sorted out according to the expected scale structure framework. After two rounds of expert consultation and a small range of test dressing, the initial scale was formed. Through the investigation of 85 patients with diabetic foot, item differentiation analysis, correlation analysis and exploratory factor analysis were used to screen the items. Cronbach’s α coefficient, retest reliability and content and structure validity analysis were used to determine the feasibility and validity of the scale.

Results: The final scale included 4 first-level items and 22 second-level items. The critical ratio method showed that the scores of each item in the high group and the low group were significantly different (P < 0.05). Correlation analysis showed that the correlation coefficient between each item and the total score was 0.431 to 0.829; The content validity index of the scale was 0.91, the exploratory factor analysis identified three common factors, and the cumulative variance contribution rate was 75.381%. The confirmatory factor analysis showed that the model fit well. The Cronbach’s α coefficient of the scale was 0.934 and the retest reliability coefficient was 0.926.

Conclusion: The self-reported outcome scale for diabetic foot patients has good reliability and validity, and can be used to investigate the health status of diabetic foot patients and evaluate the therapeutic effect.

Keywords: diabetic foot, patient-reported outcome, scale, reliability, validity

Introduction
Diabetic foot (DF) is one of the main chronic complications of diabetes, and its incidence increases with the extension of the course of diabetes. 15–25% of diabetic patients may develop DF. The prolonged course of DF has a great impact on patients’ physical and mental health. Therefore, for the symptoms of DF, unlike some acute disease symptoms, only the cure and improvement rate are used to evaluate the clinical therapeutic effect.1–5 Patients’ subjective feelings, such as quality of life and patient-reported outcomes (PRO) scale, should be increased to comprehensively evaluate the efficacy of chronic disease prevention and treatment. PRO is an assessment of their own health status without being influenced by doctors or anyone else.6–10 It is a self-observation and evaluation of the disease outcome report from the perspective of patients.

The PRO scale can be used to evaluate the health status of patients and the effect of medical intervention on diseases. It is considered as a standard method to evaluate subjective events such as symptoms. Ortega-Avila et al identified the PROMs specific for patients with DM affecting the foot and ankle and to evaluate the psychometric properties and methodological quality of these instruments. Only 11 were finally included in the qualitative review. The study drew the conclusion that criteria showed that the Foot Health Status Questionnaire (FHSQ) presented the greatest number of positive values and was the highest-quality PROM currently available for patients with DF.11 But,
with respect to foot and ankle pathologies in patients with DM, the overall methodological quality of the PROMs was low. At present, the existing special scale on the special scale for DF cannot meet clinical needs. Therefore, it is necessary to develop a PRO scale of DF that conforms to Chinese culture and is suitable for the evaluation of the clinical efficacy, as to fully reflect the feelings of patients on the treatment and nursing of the disease. Therefore, this study aimed to develop PRO scale of DF based on patients’ subjective feelings, through literature research and patient interviews in order to provide reference for accurate and comprehensive evaluation of diabetic foot patients’ condition changes. We hope that the development of this scale can provide a basis for accurately and comprehensively evaluating the quality of life for patients with DF.

Methods

Patient Selection

The patients were chosen according to the PICO structure: (P) patients were diagnosed with DF; (I) interventions were not restricted; (C) comparisons were not restricted; (O) outcomes were the PRO scales of DF. Exclusion criteria included: (C) patients with malignant metastatic tumors; (ii) patients with sepsis; (iii) patients cannot communicate properly.

Construction of Table Item Pool

A research team was established, consisting of 3 nurses and 3 clinicians. According to the PRO conceptual framework, the researcher determined the outline of the patient interview. Through one-to-one interviews with 30 patients and literature review, the interview items were sorted out to form an entry pool. The factors that affected self-quantification of patients were recorded and the relevant items were extracted using the semi-structured interview. The number of semi-structured interviews was saturated. The interview was stopped after getting the relevant data. The research team discussed and screened the entry pool and determined 24 primary items. The scale item pool of DF patients included 5 items in physiological domain, 7 items in psychological domain, 6 items in social domain and 6 items in therapeutic domain. Each item was rated by the patient according to their own understanding, on a scale of 5 to 1 from “very important” to “completely unnecessary” based on Likert grade scoring. The study was approved by the Ethics Committee of Longhua Hospital, and all patients signed a written informed consent. This study consists of development stage and validation stage. The processes were performed according to the guidelines recommended by Beaton et al.12,13

Pre-Test of Entry Pool

From Aug 2020 to Jan 2021, 45 patients with DF were randomly selected for pre-test, the contents and language expression of some items were adjusted and modified, and a total of 24 items were retained, including 5 items in the physiological domain, 7 items in the psychological domain, 6 items in social domain and 6 items in treatment domain. The first round of expert consultation questionnaire was constructed according to this item pool.

Determination of Expert Consultation Questionnaire

The expert consultation questionnaire consists of three parts: the first part is the basic information questionnaire; the second part is the current status questionnaire of PRO scale for DF patients; the third part is the self-evaluation questionnaire.

Expert Consultation

The questionnaires were submitted after completion. Further, 10 experts were asked to score the relevance and importance of each item on a scale of 1 (extremely irrelevant) to 4 (extremely relevant) and 1 (extremely unimportant) to 5 (extremely important). According to the consulting results, the items unrelated to the disease were removed and the items involving multiple aspects were split so that each item involved only one aspect. In addition, the experts modified and/or increased or decreased each item according to their professional knowledge and work experience on the scale. The preliminary scale was modified to 22 items after two rounds of Delphi surveys.

https://doi.org/10.2147/DMSO.S419841
Pre-Survey of PRO Scale
According to the method of convenient sampling, 45 patients with DF were selected to conduct a pre-survey on the contents of the pre-survey scale from Feb 2021 to Apr 2021, adjust and modify the contents and language expression of some items, determine the time needed for the scale survey, standardize the guidance paradigm, and train the investigators to unify the guidance and survey methods. Finally, 22 items were retained, including 5 items in the physiological domain, 5 items in the psychological domain, 6 items in the social domain, and 6 items in the therapeutic domain.

Reliability and Validity Test of the PRO Scale
The minimum sample size is five times the number of items of the scale. A total of 85 diabetic foot patients from 4 hospitals were selected for investigation. Inclusion criteria was: (1) meeting the diagnostic criteria for DF, (2) thinking and language expression ability is normal and (3) volunteer to participate in this study. Exclusion criteria was: (1) with various severe organ function impairment of heart, liver, kidney or brain, (2) consciousness disorder. 85 questionnaires were issued, and 85 were effectively received, with a recovery rate of 100%. According to the questionnaire number, the data were randomly divided into two groups by 1:2. Group 1 (n=28) was used for exploratory factor analysis, and group 2 (n=57) was used for confirmatory factor analysis.

Statistical Methods
Statistical software packages SPSS 22.0 and AMOS 22.0 were used for analysis. General data are described by percentage, mean and standard difference. Item discrimination analysis was used to screen the items, and the items were divided into high group (the top 27% of the total score of the scale) and low group (the bottom 27% of the total score of the scale) according to the total score of the scale. The critical ratio method was used to conduct independent test for the average score of each item in the high group and the low group, and the items with no statistical significance (P > 0.05) were deleted. Correlation analysis was performed to calculate the correlation coefficient between the score of each item in the scale and the total score of the scale. Entries with correlation coefficient ≥ 0.4 were retained. Principal component analysis was used to conduct exploratory factor analysis, and common factors with Kaiser eigenvalue > 1 were extracted. Direct rotation of the data was carried out, and the common factor with the largest contribution rate in each dimension was retained, and factor analysis was continued for the retained items. Deletion criteria: factor load < 0.40; Each common factor contains less than 3 entries; Confirmatory factor analysis was performed by maximum likelihood ratio method. Reliability test: Calculate the Cronbach’s α coefficient of the total scale and subscale, and its coefficient should be above 0.80.

Results
Expert Consultation Results
In this study, 15 experts from all over the country were selected and sent out two rounds of questionnaires. The effective recovery rate of the first round was 100%, and that of the second round was 93.3%, indicating that the experts were highly positive. The authority coefficient of experts is 0.88, indicating that experts have a high degree of authority and the research results are reliable. The agreement degree of experts in the first round was 0.78, and that in the second round was 0.86, both of which were > 0.70, suggesting that the coordination degree of experts’ opinions was good. After two rounds of expert consultation, a scale with 4 first-level items and 22 second-level items (Table 1).

Analysis of Scale Items
Item analysis of the scale. The determination values of each item ranged from 5.631 to 16.427 (P < 0.05), indicating good discrimination of each item. The correlation coefficient between each item and the total score ranges from 0.541 to 0.819 (P < 0.05). The correlation coefficient between the correction item and the total score ranges from 0.431 to 0.829 (P < 0.05), indicating that the homogeneity between the item and the whole scale is high. The correlation coefficients were all greater than 0.4, and all 22 entries were retained for further examination.
Content Validity
The content validity of the scale was evaluated according to the content validity index (CVI), the scale-level CVI (S-CVI) and the item-level CVI (I-CVI). According to the results of the second round of expert consultation, the S-CVI with 22 items was 0.91, and the I-CVI was 0.93 to 0.98.

Structural Validity
Exploratory Factor Analysis
Exploratory factor analysis was performed on Group 1. Principal component analysis was used to extract the factor with characteristic root > 1. Exploratory factor analysis was performed on the data through direct rotation. P < 0.001 was statistically significant, and the data were suitable for factor analysis. The results of exploratory factor analysis showed that the load values of all items on the common factors were all > 0.4, and the cumulative variance contribution rate was 75.381%. The three factors cover all four domains. The above items and their dimensions match the structure of the questionnaire, indicating high reliability of the structure validity (Table 2).

Confirmatory Factor Analysis
Confirmatory factor analysis was performed using data from 270 patients, and maximum likelihood estimation (MLE) was used to fit the model. The results of confirmatory factor analysis are shown in Table 3. All fitting indexes of the 3 factor models reach reference values, indicating that the model has a good degree of fitting (Table 3).

Reliability
The total Cronbach’s α coefficient of the scale was 0.934, and the Cronbach’s α coefficient of each first-level item was 0.951, 0.925, 0.881, 0.903, respectively. In this study, 135 patients in group 1 were repeatedly measured at an interval of 1 week. The scale retest reliability coefficient was 0.926, and the retest reliability coefficients of each first-level item were 0.942, 0.856, 0.878, and 0.936, respectively.
Discussion

The structural validity of the scale was evaluated by exploratory factor analysis and confirmatory factor analysis. Firstly, exploratory factor analysis was used to determine the theoretical model of the scale. In order to check whether the factor number and factor loading of the observed variables are consistent with the expectation of the pre-established theory, confirmatory factor analysis is used to further verify the correctness of the model according to the results of exploratory factor analysis. The results showed that exploratory factor analysis precipitated 3 common factors, and the loading values of 22 items on the common factors were all > 0.4, and the cumulative variance contribution rate was 75.381%. Confirmatory factor analysis showed that the model fit was good, indicating that the scale had good structural validity.

In this study, the expert consultation method was used to evaluate the content validity, and the correlation between the scale items and the dimensions was evaluated by the expert score. The results showed that the I-CVI of the scale was > 0.93, and the S-CVI was 0.91, indicating that the content validity of the scale was good.

Cronbach’s α coefficient and retest reliability were used to evaluate the reliability of this scale. Cronbach’s α coefficient is the most common used reliability coefficient, which is used to evaluate the inner homogeneity. In practical application, the reliability coefficient of the questionnaire or questionnaire should be above 0.80. The Cronbach’s α coefficient of the total volume table of this study was 0.944, and the Cronbach’s α series scores of the four first-level items of physiology, psychology, society and therapy were, 0.913, 0.935, 0.961 and 0.891, indicating that the scale had good homogeneity. Retest reliability is used to check the consistency of the results obtained by using the same measurement method at different times, also known as stability reliability. After the first measurement, the same question paper is used for the second measurement of the same

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 02</td>
<td>0.823</td>
<td>-0.234</td>
<td>0.002</td>
</tr>
<tr>
<td>Item 10</td>
<td>0.816</td>
<td>0.030</td>
<td>-0.321</td>
</tr>
<tr>
<td>Item 03</td>
<td>0.702</td>
<td>-0.111</td>
<td>0.309</td>
</tr>
<tr>
<td>Item 05</td>
<td>0.697</td>
<td>0.130</td>
<td>-0.234</td>
</tr>
<tr>
<td>Item 11</td>
<td>0.684</td>
<td>-0.211</td>
<td>0.031</td>
</tr>
<tr>
<td>Item 01</td>
<td>0.573</td>
<td>0.233</td>
<td>-0.101</td>
</tr>
<tr>
<td>Item 19</td>
<td>0.565</td>
<td>0.178</td>
<td>-0.334</td>
</tr>
<tr>
<td>Item 18</td>
<td>0.554</td>
<td>-0.311</td>
<td>0.190</td>
</tr>
<tr>
<td>Item 07</td>
<td>0.446</td>
<td>0.135</td>
<td>-0.211</td>
</tr>
<tr>
<td>Item 13</td>
<td>0.433</td>
<td>0.272</td>
<td>0.323</td>
</tr>
<tr>
<td>Item 04</td>
<td>0.399</td>
<td>0.562</td>
<td>0.278</td>
</tr>
<tr>
<td>Item 21</td>
<td>0.383</td>
<td>0.532</td>
<td>-0.331</td>
</tr>
<tr>
<td>Item 06</td>
<td>0.378</td>
<td>0.512</td>
<td>0.235</td>
</tr>
<tr>
<td>Item 16</td>
<td>0.372</td>
<td>0.498</td>
<td>-0.234</td>
</tr>
<tr>
<td>Item 22</td>
<td>0.356</td>
<td>0.472</td>
<td>0.030</td>
</tr>
<tr>
<td>Item 20</td>
<td>0.354</td>
<td>0.440</td>
<td>-0.111</td>
</tr>
<tr>
<td>Item 17</td>
<td>0.349</td>
<td>0.432</td>
<td>0.123</td>
</tr>
<tr>
<td>Item 08</td>
<td>0.342</td>
<td>0.332</td>
<td>0.534</td>
</tr>
<tr>
<td>Item 14</td>
<td>0.323</td>
<td>0.312</td>
<td>0.476</td>
</tr>
<tr>
<td>Item 09</td>
<td>0.317</td>
<td>0.308</td>
<td>0.423</td>
</tr>
<tr>
<td>Item 15</td>
<td>0.311</td>
<td>0.272</td>
<td>0.419</td>
</tr>
<tr>
<td>Item 12</td>
<td>0.309</td>
<td>0.240</td>
<td>0.403</td>
</tr>
</tbody>
</table>

Table 3 Results of Confirmatory Factor Analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>x^2/df</th>
<th>GFI</th>
<th>AGFI</th>
<th>RMSEA</th>
<th>NFI</th>
<th>CFI</th>
<th>IFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model of factors</td>
<td>2.813</td>
<td>0.921</td>
<td>0.909</td>
<td>0.071</td>
<td>0.903</td>
<td>0.918</td>
<td>0.921</td>
</tr>
<tr>
<td>Reference value</td>
<td>≤3</td>
<td>&gt;0.9</td>
<td>&gt;0.9</td>
<td>≤0.08</td>
<td>&gt;0.9</td>
<td>&gt;0.9</td>
<td>&gt;0.9</td>
</tr>
</tbody>
</table>
research object, and the correlation coefficient of each first-level item and the total score of the questionnaire is calculated. If the consistency of the two measurement results is good, then the reliability of the questionnaire retest is good. The time interval of two scales distribution depends on the specific situation of the study, usually about 1 week, the longer the time, the lower the stability coefficient is, the general requirement should be above 0.7. With the same questionnaire, the same group of respondents were tested repeatedly at a certain interval under the same circumstances as possible, and the product moment correlation coefficient of the two test results was calculated to evaluate the stability reliability of the scale. Internal consistency coefficient > 0.8, stability coefficient > 0.5 is acceptable range. In this study, all patients were repeatedly measured at an interval of 1 week, and the retest reliability coefficient of the scale was 0.926, and the retest reliability coefficient of each first-level item was 0.952, 0.856, 0.878, 0.936, respectively, indicating that the scale had a good retest reliability.

In order to ensure the progress of the study, a research team including endocrinologists and nurses specializing in diabetes was established before the study.7,8,15–18 Through great number of literatures, patient interviews and communication with experts, and two rounds of strict expert consultation, the scale was repeatedly modified according to expert advice and discussion by the research team. The initial scale finally formed covered the feelings of patients, the opinions of clinical experts and the contents of the literature, suggesting that the research process was scientific and rigorous. In the process of item selection, this study not only considered the results of statistical analysis, but also carefully considered the meaning of dimensions and the contents of items. Item differentiation analysis, correlation analysis, exploratory factor analysis, confirmatory factor analysis and reliability test were used to select items of the scale. As there is no survey scale for diabetic foot PRO in China, this scale can objectively and comprehensively evaluate the status quo of diabetic foot patients from the perspective of patients, which is conducive to providing reference for the evaluation of diabetic foot patients’ health status and observation of therapeutic efficacy.

This study also has the following deficiencies. Firstly, the patients were all from China, and its global clinical efficacy still needs to be studied. Due to the limitation of human and material resources, this scale has not been applied in clinical practice and compared with other scales. Third, there are many items in the scale and the content is relatively lengthy, which will be further refined in future research.

Conclusion
The PRO scale developed in this study has good reliability and validity, and can be used to evaluate the condition and therapeutic effect of DF patients.

Abbreviation
DF, diabetic foot.

Data Sharing Statement
All supporting data can be provided upon request to the authors.

Consent for Publication
The research was consistent with the Helsinki Declaration.

Acknowledgments
We would like to thank the reviewers for their thorough review of our manuscript, especially under the severe circumstance of worldwide epidemic COVID-19, and we wish that everybody pulls through safe and sound.

Disclosure
The authors declare that they have no competing interests in this work.

References


