

ORIGINAL REPORT

DEVELOPMENT AND REPRODUCIBILITY OF A SHORT QUESTIONNAIRE TO MEASURE USE AND USABILITY OF CUSTOM-MADE ORTHOPAEDIC SHOES

Jaap J. van Netten, MSc^{1,2}, Juha M. Hijmans, PhD^{1,3}, Michiel J.A. Jannink, PhD^{4,5},
Jan H.B. Geertzen, MD, PhD^{1,2} and Klaas Postema, MD, PhD^{1,2,3}

From the ¹Department of Rehabilitation Medicine, Center for Rehabilitation, University Medical Center Groningen, ²Graduate School for Health Research, ³School of Behavioral and Cognitive Neurosciences, University of Groningen, Groningen, ⁴Roessingh Research and Development and ⁵University of Twente, Department of Biomechanical Engineering, Enschede, The Netherlands

Objective: To develop a short and easy to use questionnaire to measure use and usability of custom-made orthopaedic shoes, and to investigate its reproducibility.

Design: Development of the questionnaire (Monitor Orthopaedic Shoes) was based on a literature search, expert interviews, 2 expert meetings, and exploration and testing of reproducibility. The questionnaire comprises 2 parts: a pre part, measuring expectations; and a post part, measuring experiences.

Patients: The pre part of the final version was completed twice by 37 first-time users before delivery of their orthopaedic shoes. The post part of the final version was completed twice by 39 first-time users who had worn their orthopaedic shoes for 2–4 months.

Results: High reproducibility scores (Cohen's kappa >0.60 or intra class correlation >0.70) were found in all but one question of both parts of the final version of the Monitor Orthopaedic Shoes questionnaire. The smallest real difference on a visual analogue scale (100 mm) ranged from 21 to 50 mm. It took patients approximately 15 min to complete one part.

Conclusion: Monitor Orthopaedic Shoes is a practical and reproducible questionnaire that can measure relevant aspects of use and usability of orthopaedic shoes from a patient's perspective.

Key words: orthopaedic shoes, usability, questionnaire, patient satisfaction.

J Rehabil Med 2009; 41: 913–918

Correspondence address: Jaap J. van Netten, Center for Rehabilitation, University Medical Center Groningen, PO Box 30.001, NL-9700 RB Groningen, The Netherlands. E-mail: j.j.van.netten@rev.umcg.nl

Submitted March 3, 2009; accepted June 17, 2009

INTRODUCTION

Custom-made orthopaedic shoes (OS) can be used to diminish or prevent serious foot and/or ankle problems for a wide-range of patient groups (e.g. patients with diabetes, rheumatoid arthritis, degenerative foot disorders, muscle diseases) (1–3). OS are frequently prescribed in the Netherlands, with almost

50,000 pairs provided in 2006 (4). The total cost of OS in that year was almost 60 million Euros (4). In England and Wales, approximately 200,000 persons were provided with OS by the National Health Service (NHS) in 2000, with an expenditure of almost 40 million pounds sterling (5).

The non-use of OS is a well-known problem. Non-use rates vary, ranging from 20% to 25% for first-time users (2, 6, 7), and from 4% to 19% when experienced users who were provided with a subsequent pair of OS were also taken into account (8–11). Non-use is influenced by the usability of OS. Usability is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction, in a specified context of use” (International Organization for Standardization (ISO), 9241-11). Within the domains of usability stated, effectiveness, efficiency and satisfaction, the following aspects are associated with non-use of OS: benefits of OS with regard to walking capacities, wound healing, pain, etc. (domain effectiveness); comfort and ease of use, and the efficiency of the delivery process of OS (domain efficiency); communication with prescribing specialists, and cosmetic appearance (domain satisfaction) (2, 6–21). However, in a systematic review it was concluded that none of the randomized clinical trials focused on all these domains of usability (22). Subsequently, Jannink et al. (2) focused on all domains of usability, however, only in patients with degenerative disorders of the foot. Therefore, further insight into the usability of OS in a wide range of patient groups is required.

Of the instruments that are available to gain insight into aspects of use and usability of OS, only the Questionnaire for Usability Evaluation (QUE) quantifies aspects of all domains of usability as defined by the ISO (22, 23). However, the QUE was developed specifically for patients with degenerative disorders of the foot. Even more importantly, the QUE is a lengthy and time-consuming questionnaire (it takes about 30 min to complete one questionnaire (pre or post part)), which limits its practical application. Therefore, the aim of the present study was to develop a short and easy to use questionnaire that could measure the most relevant aspects of use and usability of OS from a patient's perspective, in a wide-range of patient groups. In addition, reproducibility of the questionnaire was studied.

METHODS

The development of the questionnaire, Monitor Orthopaedic Shoes (MOS), comprised 7 phases: 1) literature search and expert interviews; 2) development of the first (pilot) version (termed MOS^{v1}); 3) expert meetings to discuss MOS^{v1}; 4) development of the second version (MOS^{v2}); 5) exploration of the reproducibility of MOS^{v2}; 6) development of the final version (MOS^{fv}); and 7) testing MOS^{fv} for reproducibility. The procedures and patient inclusion criteria were the same for exploration (phase 5) and testing (phase 7) of reproducibility, and will therefore be described in the paragraph concerning phase 5. The methodology of the development of MOS is shown in Fig. 1.

MOS was developed specifically for first-time users of OS, because there is a large difference between first-time and experienced users, especially with regard to their expectations (8). MOS was developed primarily in Dutch, and reproducibility was defined in the Dutch language.

Phase 1 – Literature search and expert interviews

Articles were sought in MEDLINE (1989–2008) and EMBASE (1989–2008), using combinations of the following keywords: orthopaedic, therapeutic, surgical, prescribed, shoe, foot, patient satisfaction, usability, non-use, diabetes mellitus, rheumatic diseases, foot deformities, foot-diseases. The combination of keywords used in MEDLINE is shown in Appendix I. This combination was adapted to suit EMBASE. The initial selection of articles was based on the title and the content of the abstract. The reference lists of relevant publications found were checked carefully. A total of 17 articles concerning use and usability of OS were identified (2, 6–21).

To obtain additional information from clinical practice, semi-structured expert interviews were held with specialists in rehabilitation medicine ($n=3$), certified orthopaedic shoe technicians ($n=3$) and an experienced user of OS ($n=1$). These experts were interviewed about the importance of the different aspects of usability of OS, different goals of OS, the relevance of different activities of the user, and the expectations of users.

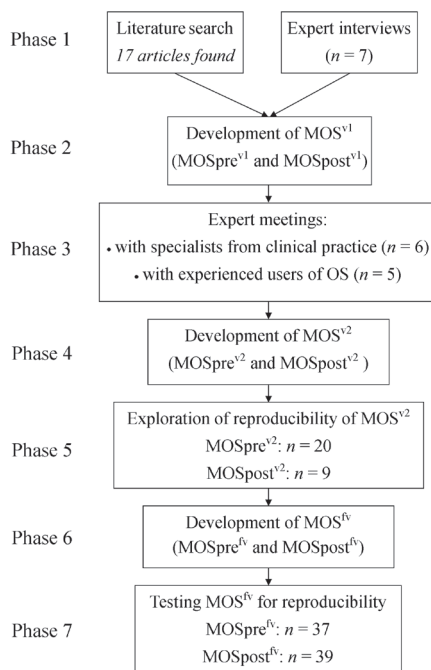


Fig. 1. Methodology of the development of the Monitor Orthopaedic Shoes (MOS) questionnaire. OS: orthopaedic shoes.

Phase 2 – Development of MOS^{v1}

Based on the literature search and the expert interviews, the most relevant aspects of use and usability of OS, divided into the domains of usability, were determined and operationalized into the questions of MOS^{v1}. MOS^{v1} comprised 2 parts, a pre part (MOSpre^{v1}) and a post part (MOSpost^{v1}). MOSpre^{v1} is completed by patients prior to receiving their first pair of OS. MOSpre^{v1} was designed to measure the expectations of the patient and the current status of several outcome variables. MOSpost^{v1} was designed as a follow-up to measure experiences with use and usability of patients' OS and the change in outcome variables. MOSpost^{v1} is completed after patients have worn their OS for 2–4 months, with no insight into the answers of MOSpre^{v1}. The interviewed experts indicated that a period of 2 months is necessary to become accustomed to OS and that after 4 months problems with wearing-out of OS may have already developed.

MOS^{v1} comprised questions in different forms: multiple choice, visual analogue scale (VAS), open, and photo-based questions.

Phase 3 – Expert meetings

Content validity was taken into account during all phases of the development of MOS. It was not possible to determine criterion validity, because no similar measuring instrument to MOS exists. The QUE, which is most closely related, was developed only for patients with degenerative disorders of the foot, whereas MOS was developed for a wide-range of patient groups.

In order to ensure good content validity, 2 meetings were organized with experts who did not take part in a previous phase of this study. One meeting was held with specialists in rehabilitation medicine ($n=2$) and certified orthopaedic shoe technicians ($n=4$), and a second with experienced users of OS ($n=5$). In these meetings, all experts were asked to comment on all aspects of MOS^{v1}. A general discussion was held during both meetings regarding the relevance of the questions, and any important aspects missing about the use and usability of OS. Following this, a specific discussion was held regarding the clarity and practical application of each question in MOS^{v1}.

Phase 4 – Development of MOS^{v2}

Based on the recommendations and suggestions made during the expert meetings, MOS^{v2} was developed. MOS^{v2} was sent to the participants of both expert meetings again for final comment. This did not result in any changes.

As MOS^{v1}, MOS^{v2} comprised 2 parts: a pre part, measuring expectations and the current situation (MOSpre^{v2}), and a post part, measuring experiences and changes (MOSpost^{v2}). MOS^{v2} comprised questions in different forms: multiple choice, VAS, open, and photo-based questions. The expert meetings and pilot-testing indicated that patients understood the direction of the choices and how to answer the questions.

Phase 5 – Exploration of reproducibility of MOS^{v2}

Patients. Inclusion criteria for patients were: (i) first-time user of OS; (ii) 16 years of age or older; (iii) able to read Dutch; and (iv) able to complete the questionnaire without help related to cognitive or physical impairments.

Study design. Although MOS^{v2} comprised 2 related parts, MOSpre^{v2} and MOSpost^{v2}, for practical reasons reproducibility was explored in 2 separate groups in a within-group design. Patients were recruited from 2 participating orthopaedic shoe companies (OIM Orthopedie, Haren, The Netherlands and Roessingh Revalidatie Techniek (RRT), Enschede, The Netherlands). For MOSpre^{v2}, patients meeting the inclusion criteria were recruited by the orthopaedic shoe technician after their visit for foot measurements for a first pair of OS, whereas for MOSpost^{v2} patients were telephoned by the orthopaedic shoe technician if they had worn their first pair of OS for 2–4 months. After patients gave informed consent, MOS^{v2} was sent to them by post. Once MOS^{v2} was completed for the first time and received by the researchers,

MOS^{v2} was again sent to patients by post. Because of privacy matters, researchers had no insight into the personal characteristics of patients who did not give informed consent. The procedures of this study were approved by the medical ethics committee of the University Medical Center Groningen.

Data analysis. Reproducibility was analysed using SPSS for Windows, version 16.0 (SPSS Inc. Chicago, IL, USA). The reproducibility of the multiple choice questions was analysed using Cohen's kappa. A kappa value of less than 0.40 indicates poor to fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1.00 almost perfect agreement (24). The reproducibility of the VAS questions was analysed using an intra-class correlation (ICC), absolute agreement, 2-way random model for single measures. This model measures the degree of absolute agreement among the scores on the first and the second completed questionnaire (25). In general, values over 0.70 are considered as high agreement (24). The 95% confidence interval of the ICC was also calculated. The open and photo-based questions were not included in the reproducibility analysis.

In addition, the standard error of measurement (SEM) and the smallest real difference (SRD) of the VAS questions were also calculated. The SEM is the square-root of the within-subject variance (26), the SRD is $1.96 \times \sqrt{2} \times \text{SEM}$ (27). Both are expressed in the same dimension as the measurement. The SEM provides an interpretation of the magnitude of within-subject variability, whereas the SRD indicates the point where the difference between 2 consecutive measurements within a subject can be interpreted as a real difference (26, 27).

Results of the exploration of reproducibility. MOSpre^{v2} was completed twice by 20 out of 23 patients; MOSpost^{v2} was completed twice by 9 out of 13 patients. The interval between completing MOS^{v2} for the first and the second time varied between 1 and 3 weeks. This is a usual test-retest interval (26). It was not expected that any clinically relevant changes in patients' health status occurred in this period. High reproducibility scores (>0.70) were found in 47 out of 62 questions. The SRD showed a range of 18–50 mm. Because this was only an exploration of the reproducibility in a small group of patients, results will not be presented in detail.

Additional questions were answered by all patients regarding: the inclusion of all relevant aspects of use and usability of OS; whether each question was clear to them; and different aspects of clarity and practical application of MOS^{v2} (for example: "Did you have enough space for your answers?"; "Is the font clear and readable?"; "What do you think of the length of MOS^{v2}?"). Six questions (3 in MOSpre^{v2} and 3 in MOSpost^{v2}) were not clear for the patients. Patients were satisfied with the aspects of clarity and practical usefulness.

Phase 6 – Development of MOS^{fv}

Based on the results of this exploration, MOS^{fv} was developed. Compared with MOS^{v2}, 8 questions were deleted (5 in MOSpre^{v2} and 3 in MOSpost^{v2}), 11 questions were adjusted (7 in MOSpre^{v2} and 4 in MOSpost^{v2}), and 4 questions were added (2 in both MOSpre^{fv} and MOSpost^{fv}). It was ensured that the most relevant aspects of use and usability were still covered in MOS^{fv}.

As with the previous versions, MOS^{fv} comprised 2 parts: MOSpre^{fv}, measuring expectations and the current situation by means of 15 multiple choice questions and 12 VAS questions, and MOSpost^{fv}, measuring experiences and changes by means of 11 multiple choice questions and 19 VAS questions. Both parts also contain open (5 and 7, respectively) and photo-based (2 in both) questions. MOS^{fv} was sent to the participants of both expert meetings again for comments. This did not result in any changes, and experts indicated that the most relevant aspects of use and usability were included in MOS^{fv}.

Phase 7 – Testing MOS^{fv} for reproducibility

The same inclusion criteria, procedures, and data analysis as in phase 5 were applied, apart from 2 differences with regard to the recruit-

ment of patients. Patients were recruited from OIM Orthopedie and Penders Voetzorg (Heythuysen, The Netherlands), instead of OIM Orthopedie and RRT, because Penders Voetzorg is a larger company with more patients. Furthermore, the same method of administration was of application for MOSpre^{fv} and MOSpost^{fv}: at the same time, all patients from the 2 participating orthopaedic shoe companies meeting the inclusion criteria were contacted by post by the orthopaedic shoe company. For MOSpre^{fv}, a mailing was sent to patients who had visited the orthopaedic shoe company for foot measurements for a first pair of OS in the last 2 months; for MOSpost^{fv}, a mailing was sent to patients who had worn their first pair of OS for 2–4 months. The mailing contained an information letter and an informed consent form, the letter could be returned to the researchers by pre-stamped envelope. After receiving a consent form, MOS^{fv} was sent to patients by post. Once MOS^{fv} was completed for the first time and received by the researchers, MOS^{fv} was again sent to patients by post.

MOSpre^{fv} was sent to 58 patients. 37 patients completed MOSpre^{fv} twice; 3 patients completed MOSpre^{fv} only once. MOSpost^{fv} was sent to 49 patients. 39 patients completed MOSpost^{fv} twice; 2 patients completed MOSpost^{fv} only once. The interval between completing MOS^{fv} for the first and the second time varied between 1 and 3 weeks. This is a usual test-retest interval (26). It was not expected that any clinically relevant changes in patients' health status took place in this period. Patient characteristics are summarized in Table I. There were no major demographic differences between the groups.

RESULTS

Tables II and III list the results with respect to the reproducibility analysis, for MOSpre^{fv} and MOSpost^{fv}, respectively. Cohen's kappa for 13 questions was between 0.61 and 0.80, indicating substantial agreement. Cohen's kappa for 8 questions was between 0.81 and 1.00, indicating almost perfect agreement. Cohen's kappa of one question (regarding the patient's expectation of other people's opinion of the cosmetic appearance) was 0.47, which indicates moderate agreement. Detailed analysis of this question revealed that most variation was between the response possibilities "neutral" and "beautiful" or "neutral" and "I do not know". The ICC of all questions was above 0.70, which indicates high agreement. The SRD was in the range 21–50 mm. It took patients approximately 15 min to complete either the pre or the post part.

Table I. Characteristics of patients participating in the reproducibility study of the final version of the Monitor Orthopaedic Shoes (MOS^{fv}) questionnaire

Characteristics	MOSpre ^{fv} (n=37)	MOSpost ^{fv} (n=39)
Age, years, mean (SD)	66 (13)	67 (13)
Gender, female, n (%)	26 (70)	22 (56)
Disorder, n (%)*		
Diabetes mellitus	9 (24)	13 (33)
Rheumatoid arthritis	10 (27)	7 (18)
Foot disorder†	22 (60)	29 (74)
Muscle disease	4 (11)	2 (5)
Other disease	10 (27)	12 (31)

*More than one disorder was possible; disorders were indicated by patients themselves.

†Foot disorder was unspecified.

SD: standard deviation.

Table II. Results of the reproducibility analysis of the pre part of the final version of the Monitor Orthopaedic Shoes (MOSpre^{iv}) questionnaire (n = 37)

Domain	n	Cohen's kappa	n	ICC (CI)		SRD, mm	
				Smallest	Largest	Smallest	Largest
Use	1	0.66					
Effectiveness							
Walking capacity	3	0.72–0.76					
Wounds	2	0.88–1.0					
Pain			4	0.76 (0.56–0.87)	0.86 (0.74–0.93)	29	50
Sprains			2	0.76 (0.42–0.92)	0.93 (0.87–0.97)	25	31
Efficiency							
Comfort in use			1	0.76 (0.57–0.87)		27	
Satisfaction							
Communication	4	0.72–0.80	2	0.84 (0.71–0.91)	0.84 (0.68–0.92)	26	27
Cosmetic appearance	2	0.47–1.0	1	0.72 (0.50–0.85)		42	
Acceptability			2	0.84 (0.70–0.92)	0.85 (0.71–0.92)	21	31

Note: Three questions concerning demographic information are not shown.

n: number of questions, the range of the Cohen's kappa is shown; ICC: Intra Class Correlation; CI: confidence interval; the smallest and largest ICC with corresponding CI are shown; SRD: smallest real difference; the smallest and largest SRD are shown.

DISCUSSION

The result of this study is a practical and reproducible questionnaire that can measure the most relevant aspects of use and usability of OS from a patient's perspective, for a wide-range of patient groups. This questionnaire, MOS, can be applied at 2 levels: (i) to measure expectations (prior to prescription) and experiences with use and usability of OS in a group of patients; and (ii) to identify problems with OS on an individual level, prior to provision or after experience with OS. Insight into the use and usability of OS on both levels may lead to clearer identification of problems relating to the usability of OS, an increase in patient satisfaction, and an increased rate of use of OS.

It can be concluded that MOS is a reproducible questionnaire, because high agreement was found in all but one question of the final version. This one question explored the patient's expectation of other people's opinion of the cosmetic appearance of their OS. It was apparent that patients' concern for other

people's opinion 3 months before actual delivery of their OS is subject to much variation. As the detailed analysis showed, answers in the categories "neutral", "beautiful", or "I do not know" varied the most, thus these answers should be treated with care. It may be even better to not interpret this question at all. In contrast, the same question in the experienced part of MOS resulted in "almost perfect agreement".

Relatively large SRD values were found, ranging from 21% to 50% of the VAS. When a part of MOS is completed twice by the same patient (e.g. by administering MOSpost after 3 and after 12 months of using OS), differences between consecutive scores should be large (greater than 21–50%) before they can be interpreted as a real difference. However, most scores were found near either one of the extreme end-points (e.g. patients either have a lot of pain or almost no pain; shoes fit either very badly or very well; etc.). It can be expected that a clinically relevant change will correspond with a shift from one extreme end-point to the other. The resulting large difference between the 2 scores will then likely be larger than the SRDs found.

Table III. Results of the reproducibility analysis of the post part of the final version of the Monitor Orthopaedic Shoes (MOSpost^{iv}) questionnaire (n = 39)

Domain	n	Cohen's kappa	n	ICC (CI)		SRD, mm	
				Smallest	Largest	Smallest	Largest
Use	5	0.75–0.90	1	0.92 (0.85–0.96)			
Effectiveness							
Walking capacity	2	0.78–0.89					
Wounds	2	0.93–1.0					
Pain			4	0.78 (0.52–0.91)	0.95 (0.89–0.98)	23	48
Sprains			2	0.84 (0.62–0.94)	0.93 (0.85–0.96)	22	27
Efficiency							
Comfort in use			7	0.71 (0.50–0.85)	0.91 (0.82–0.95)	33	49
Satisfaction							
Communication			2	0.76 (0.55–0.88)	0.81 (0.66–0.90)	37	41
Cosmetic appearance	1	0.82	1	0.86 (0.75–0.93)		36	
Acceptability			2	0.72 (0.52–0.85)	0.73 (0.51–0.86)	41	42

Note: One question concerning demographic information is not shown.

n: number of questions, the range of the Cohen's kappa is shown; ICC: intra class correlation; CI: confidence interval; the smallest and largest ICC with corresponding CI are shown; SRD: smallest real difference; the smallest and largest SRD are shown.

Consequently, the large SRDs do not present a problem in the application of MOS.

Validity is regarded as the most fundamental consideration in developing and evaluating a measuring instrument (28). Various categories of validity have been described in the literature. Of these, construct validity is the overarching category that can be addressed in 2 different ways, by criterion and content validity (26, 29). Criterion validity is not relevant in this study, as no similar measuring instrument exists that can be used as a criterion. Emphasis has therefore been placed on the content validity. Good content validity increases motivation and reduces dissatisfaction among users (e.g. OS technicians or medical specialists) and respondents (e.g. patients), and makes it more likely that other stakeholders (e.g. policy makers, health insurance companies) accept the results (26). These are all essential aspects when considering application of MOS. Apart from the opinion of experts, there is no method to analyse content validity. MOS is based on a literature search and all experts indicated that the most relevant aspects of use and usability of OS are included. We conclude that content validity of MOS can therefore be regarded as satisfactory.

All questions in MOS were formulated in a manner that was preferred by most experts. This is a strong point from a respondent-focused view, because it enhances clarity and the practical application of the questionnaire. On the other hand, it is also a limitation. The questions in MOS are formulated in different ways by using either a VAS or multiple choices with 2, 3, 4, or 5 response categories. Consequently, it is not possible to calculate an interpretable sum-score or to perform a factor-analysis on MOS. This means that the underlying constructs of the questionnaire cannot be determined. With MOS, it is therefore not possible to identify the relationships between the different aspects of usability or the relative importance of these aspects with regard to use. MOS can, however, be used to gain insight in patients' opinions with regard to the measured aspects of use and usability separately.

A possible limitation of the reproducibility study is that MOS_{pre}^{iv} and MOS_{post}^{iv} were not completed by the same subjects. However, reproducibility was determined through a within-group design that is not affected by differences between the groups. Furthermore, both parts of MOS^{iv} were completed by patients at appropriate timing (i.e. prior to receiving OS or after having used OS for 2–4 months). If the same patients had completed both parts, there would have been an interval of approximately 6 months between completing the pre and the post part. In that period, relevant changes in a patient's health status may occur. Finally, there are no reasons to expect different reproducibility scores between the groups, because there were no large differences in the patient characteristics of both groups and the methods of administration of the patients were the same.

With regard to the application of MOS, some remarks can be made. First, MOS was specifically designed for first-time users. The difference between first-time and experienced users is essential in the interpretation of outcomes of MOS. First-time users compare their (future) pair of OS with their normal shoes, whereas experienced users compare their new pair of OS with their previous pair(s) of OS and possibly with

previous normal shoes. This different frame of reference may influence expectations and experiences of a user's (future) pair of OS, and thereby influence the outcomes of MOS. Therefore, when MOS is administered to a group of patients, the outcomes of first-time users and experienced users should be separated. Furthermore, it should be kept in mind that some questions may need to be added when MOS is administered to experienced users. For example, with regard to the number of pairs that have previously been worn. The second remark concerns the language of the questionnaire. MOS has been developed specifically for the Dutch situation. Cross-cultural adaptation of MOS to other countries will be a useful step to enhance research regarding use and usability of OS around the world. The Dutch version of MOS and its English version (unilaterally translated by the authors) are available from the corresponding author.

In conclusion, MOS is the first practical and reproducible questionnaire that can measure the most relevant aspects of use and usability of OS from a patient's perspective, for a wide range of patient groups. MOS is short and easy to complete, and can be used for evaluation of a group of patients, as well as for assessment of an individual patient.

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support provided by the Development Fund for the Orthopaedic Shoe Company (OFOM); by the National Centre for Innovation in Rehabilitation Technology; by the Foundation OIM Orthopedie; and by the Foundation Beatrixoord North–Netherlands, The Netherlands. The authors would like to thank J. Stötefalk, G. Roelofs and I. Meijers and their orthopaedic shoe companies OIM Orthopedie (Haren, The Netherlands), RRT (Enschede, The Netherlands), and Penders Voetzorg (Heythuysen, The Netherlands), for their co-operation in this project. The Dutch version of MOS^{iv} and its unilateral translated English version can be obtained from the corresponding author.

REFERENCES

1. Boer H, Seydel ER. Medical opinions, beliefs and prescription of orthopaedic footwear: a survey of Dutch orthopaedists and rehabilitation practitioners. *Clin Rehabil* 1998; 12: 245–253.
2. Jannink MJ, IJzerman MJ, Groothuis-Oudshoorn K, Stewart RE, Groothoff JW, Lankhorst GJ. Use of orthopedic shoes in patients with degenerative disorders of the foot. *Arch Phys Med Rehabil* 2005; 86: 687–692.
3. Postema K, Toornend JLA, Zilvold G, Schaars AH. Orthopedisch maatschoeisel in de medische praktijk [Custom-made orthopaedic shoes in the medical practice]. Houten/Antwerpen: Bohn Stafleu Van Loghum; 1991 (in Dutch).
4. GIP, College voor Zorgverzekeringen. Monitor hulpmiddelen 2006 [Monitor medical devices]. [Updated 2006 Jun 22; cited 2007 Oct 24] Available from: http://www.cvz.nl/resources/rpt0606%20monitor%20hulpmiddelen_tcm28-21637.pdf (in Dutch).
5. Audit Commission. Fully equipped: the provision to older people or disabled people by the NHS and social services in England and Wales. London: Audit Commission; 2000.
6. Baker N, Letherdale B. Audit of special shoes: are they being worn? *Diabet Foot* 1999; 2: 100–104.
7. Costigan PS, Miller G, Elliott C, Wallace WA. Are surgical shoes providing value for money. *BMJ* 1989; 299: 950.
8. Dijkstra B, de Witte LP, Gelderblom GJ, Wessels RD, Soede M.

- Non-use of assistive technology in the Netherlands: a non-issue? *Disabil Rehabil Assist Technol* 2006; 1: 97–102.
9. Fisher LR, McLellan DL. Questionnaire assessment of patient satisfaction with lower-limb orthoses from a district hospital. *Prosthet Orthot Int* 1989; 13: 29–35.
 10. Knowles EA, Boulton AJ. Do people with diabetes wear their prescribed footwear? *Diabet Med* 1996; 13: 1064–1068.
 11. Philipsen AB, Ellitsgaard N, Krosgaard MR, Sonne-Holm S. Patient compliance and effect of orthopaedic shoes. *Prosthet Orthot Int* 1999; 23: 59–62.
 12. Breuer U. Diabetic patient's compliance with bespoke footwear after healing of neuropathic foot ulcers. *Diabetes Metab* 1994; 20: 415–419.
 13. Cavanagh PR. Therapeutic footwear for people with diabetes. *Diabetes Metab Res Rev* 2004; 20 Suppl 1: S51–S55.
 14. Herold DC, Palmer RG. Questionnaire study of the use of surgical shoes prescribed in a rheumatology outpatient clinic. *J Rheumatol* 1992; 19: 1542–1544.
 15. Johnson M, Newton P, Goyder E. Patient and professional perspectives on prescribed therapeutic footwear for people with diabetes: a vignette study. *Patient Educ Couns* 2006; 64: 167–172.
 16. Macfarlane DJ, Jensen JL. Factors in diabetic footwear compliance. *J Am Podiatr Med Assoc* 2003; 93: 485–491.
 17. Newson F, Robinson N, Smith S. Orthopaedic footwear: a quality initiative to find out what patients think. *Physiotherapy* 1992; 78: 12–14.
 18. Park C, Craxford AD. Surgical footwear in rheumatoid-arthritis – a patient acceptability study. *Prosthet Orthot Int* 1981; 5: 33–36.
 19. Van der Esch M, Heijmans M, Dekker J. Factors contributing to possession and use of walking aids among persons with rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 2003; 49: 838–842.
 20. Williams A, Meacher K. Shoes in the cupboard: the fate of prescribed footwear? *Prosthet Orthot Int* 2001; 25: 53–59.
 21. Williams AE, Nester CJ, Ravey MI. Rheumatoid arthritis patients' experiences of wearing therapeutic footwear – a qualitative investigation. *BMC Musculoskelet Disord* 2007; 8: 104.
 22. Jannink MJ, van Dijk H, de Vries J, Groothoff JW, Lankhorst GJ. A systematic review of the methodological quality and extent to which evaluation studies measure the usability of orthopaedic shoes. *Clin Rehabil* 2004; 18: 15–26.
 23. Jannink MJ, de Vries J, Stewart RE, Groothoff JW, Lankhorst GJ. Questionnaire for usability evaluation of orthopaedic shoes: construction and reliability in patients with degenerative disorders of the foot. *J Rehabil Med* 2004; 36: 242–248.
 24. Landis JR, Koch GG. Measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
 25. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996; 1: 30–46.
 26. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 3rd edn. Oxford: Oxford University Press; 2003.
 27. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* 2001; 10: 571–578.
 28. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association; 1999.
 29. Trochim WM. Measurement validity types. The Research Methods Knowledge Base [updated 2006 Oct 20; cited 2009 Feb 24]. Available from: www.socialresearchmethods.net/kb/measval.php.

APPENDIX I. Combination of key words used in the literature search in Medline

#1	((Orthopaedic or orthopedic or therapeutic or surgical or prescrib*) near1 (shoe* (MeSH) or foot* (MeSH))
#2	Patient satisfaction or usability or non-use
#3	Diabetes mellitus (MeSH)
#4	Rheumatic diseases (MeSH)
#5	Foot deformities (MeSH)
#6	Foot diseases (MeSH)
#7	PY > 1989
#8	#1 and #2
#9	#3 or #4 or #5 or #6
#10	#8 and #9 and #7
