Development of grading scales of pedal sensory loss using Mokken scale analysis on the Rotterdam Diabetic Foot Study Test Battery data

Willem D. Rinkel MD1,2,3 | M. Hosein Aziz MD1 | Johan W. Van Neck PhD1 | Manuel Castro Cabezas MD, PhD4 | L. Andries van der Ark PhD5 | J. Henk Coert MD, PhD2,3

1Department of Plastic, Reconstructive, and Hand Surgery, Erasmus University Medical Center, Rotterdam, The Netherlands
2Department of Plastic, Reconstructive, and Hand Surgery, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands
3Department of Plastic, Reconstructive and Hand Surgery, University Medical Center Utrecht, Utrecht, The Netherlands
4Department of Internal Medicine/Centre for Diabetes, Endocrinology and Vascular Medicine, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands
5Faculty of Social and Behavioural Sciences, Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

Correspondence
Willem D. Rinkel, Department of Plastic, Reconstructive, and Hand Surgery, University Medical Center Utrecht, Utrecht, The Netherlands.
Email: w.d.rinkel@umcutrecht.nl

Abstract

Introduction: Loss of sensation due to diabetes-related neuropathy often leads to diabetic foot ulceration. Several test instruments are used to assess sensation, such as static and moving 2-point discrimination (S2PD, M2PD), monofilaments, and tuning forks.

Methods: Mokken scale analysis was applied to the Rotterdam Diabetic Foot Study data to select hierarchies of tests to construct measurement scales.

Results: We developed 39-item and 31-item scales to measure loss of sensation for research purposes and a 13-item scale for clinical practice. All instruments were strongly scalable and reliable. The 39 items can be classified into 5 hierarchically ordered core clusters: S2PD, M2PD, vibration sense, monofilaments, and prior ulcer or amputation.

Discussion: Guided by the presented scales, clinicians may better classify the grade of sensory loss in diabetic patients' feet. Thus, a more personalized approach concerning individual recommendations, intervention strategies, and patient information may be applied.

KEYWORDS
diabetic sensorimotor polyneuropathy, early detection, grading loss of sensation, medical decisionmaking, neuropathy, psychometrics, risk stratification, scale development

INTRODUCTION

Diabetic sensorimotor polyneuropathy (DSP) occurs in about 50% of diabetic patients, leading to decreased quality of life and increased mortality.1-3 Sensory loss due to DSP is one of the most important risk factors for diabetic foot ulceration and amputation.4 DSP is frequently accompanied by positive sensory phenomena. Yet, it is the simultaneous process of decreased sensation that places the feet of diabetic...
patients at risk. Because not every patient with positive sensory symptoms has sensory loss, it is useful to focus on objective measures such as foot sensation to assess large-fiber nerve function.

Loss of sensation can be assessed with several instruments. Current guidelines recommend an annual screening with a 10-g Semmes-Weinstein monofilament (SWM) or a tuning fork. These instruments test different somatosensory corpuscles and nerve fibers, functions of which are progressively lost during the natural course of neuropathy. However, a cutaneous threshold of ≥10 g is an indicator of large-fiber demyelination, which becomes informative in a late stage of neuropathy and reflects a high risk for foot ulceration and lower extremity amputation. Other measurements, such as 2-point discrimination, have proven to be early indicators of nerve pathology and may be able to detect earlier alterations in foot sensation.

Few available studies have described the sequence in which sensory tests become abnormal in the natural course of diabetes-related neuropathy. To more precisely categorize patients with diabetes according to their degree of sensory loss, we applied the Mokken scale analysis (MSA), which is related to nonparametric forms of item-response theory (IRT), to the tests used in the Rotterdam Diabetic Foot (RDF) Study. MSA is a scaling method increasingly used in health sciences. IRT assumes that a latent (not directly observable) trait, denoted as θ (theta), drives the patients’ scores on the items. Typically, θ is used as a proxy for the construct being measured (ie, foot sensation, hence foot sensation determines the item scores). Because θ is latent, a patient’s score on θ must be estimated from the observable item scores. MSA is a flexible scaling method, in contrast to other IRT models such as Rasch analysis, so it fits data relatively well and includes a fair amount of items that can be used for ordinal measurement. In the present analysis, we used questionnaire data and the results of sensory tests of the feet as items (RDF Study Test Battery).

We investigated whether the tests of the RDF Study Test Battery were unidimensional, scalable, and reliable in assessing sensation in the foot. Information on the degree of sensory loss may help clinicians to assess the risk of lower extremity complications, resulting in more personalized recommendations regarding intervention strategies and patient information.

2 METHODS

2.1 Study design and subjects

Between January 2014 and June 2015, patients were evaluated in the outpatient Diabetes Clinic of the Franciscus Gasthuis in Rotterdam, The Netherlands, as part of the RDF Study—a prospective cohort study that investigates the deterioration of sensation in diabetic patients’ feet over time. The RDF Study design and methods are described in more detail in previous studies. Inclusion criteria included patients diagnosed with diabetes mellitus (treated by oral blood glucose–lowering drugs and/or insulin), who were ≥18 years old, spoke Dutch or English, and had no significant cognitive impairment. Exclusion criteria were assessed at the interview and with a screening questionnaire and included a positive history of active radicular syndrome or a neurological disease that interfered with sensation in the feet. Demographic data were obtained from the patients’ files. All subjects provided written informed consent. The institutional review board and the medical ethics committee of Erasmus MC University Center, Rotterdam, The Netherlands, approved the study (MEC-2009-148).

2.2 Comparison with healthy controls without known neuropathy

A total of 196 healthy volunteers were tested with the same measurement instruments and the same protocol as the RDF Study population as part of a separate study to obtain normative test values. Volunteers were recruited from hospital and university personnel and relatives and friends of patients visiting the outpatient clinic. Patients were included in the study if they were ≥18 years of age, had no significant cognitive impairment, spoke Dutch or English, and provided signed informed consent. Exclusion criteria were a positive history of active radicular syndrome, a neurological disease that interfered with sensation in the feet, diabetes mellitus, thyroid malfunction, alcohol abuse, human immunodeficiency virus, or chemotherapy—all these were established at the interview using a screening questionnaire. Data sets were combined to compare patients with healthy controls.

2.3 Measurement instrument

2.3.1 Rotterdam Diabetic Foot Study Test Battery

Patients and volunteers were screened using monofilaments, static and moving 2-point discrimination tests, a tuning fork, cold sensation tests, the Michigan Neuropathy Screening Instrument (MNSI), and the Romberg test. Information on prior ulceration and amputation, as indicators of severe sensory loss, was retrieved from the patient’s file and interview. Cutaneous threshold (1-point static discrimination, S1PD) was tested on 5 locations of each foot using SWMs (Baseline; Tactile Fabrication Enterprises, White Plains, NY, USA) ranging from 0.008 to 300 g. The test locations were chosen in concordance with the nerve distribution in the foot (see Figure S1 online). Areas with excessive callus formation were avoided. Innervation density (determined by static and moving 2-point discrimination tests: S2PD and M2PD) was assessed on the same test locations using a Disk-Criminator (US Neurologicals LLC, Poulso, Washington, USA). M2PD was not assessed at the fifth toe because the area is too small to conduct the test. A Rydel-Seiffer tuning fork (Martin, Tuttlingen, Germany) tested the vibration threshold on the medial malleolus and dorsal interphalangeal joint of the hallux. Both feet were examined. Neuropathy symptoms were assessed using the MNSI questionnaire, which was administered before the physical examination.

Using the Tinel sign, we scored localized nerve compression as positive when tingling and electrical shocks were elicited after tapping the tibial nerve at the left and right tarsal tunnel. To test cold perception, a cold piece of metal was bilaterally applied to the skin of the foot arc. Proprioception was tested using the Romberg test.
2.4 | Data analysis

A cross-sectional analysis of the RDF Study baseline data was carried out for the MSA. Results from the RDF Study Test Battery (both tests of sensation and questionnaire data) were dichotomized because MSA requires that all items have the same number of categories. Individual sensory test items were comprised of both a sensory test and the test location (eg, S1PD at the hallux is labeled as S1PD I, and S2PD at the medial heel as S2PD II). Based on previously published normative values, the threshold for S2PD and M2PD was set at 8 mm and at 10 g for S1PD. Vibration threshold was compared with age-related reference values. Positive symptoms (eg, tingling and burning sensations) and negative symptoms (eg, numbness) were retrieved from the MNSI questionnaire, resulting in two items. When scoring at or below the threshold (ie, "could feel the stimulus"), 0 was noted. A score of 1 was noted when a patient scored above the threshold, meaning aberrant (ie, "could not feel the stimulus"). In total, 42 individual RDF Study Test Battery items were identified per subject. Second and third annual follow-up data are presented and compared with baseline data as a measure of the statistical significance of the change scores.

2.5 | Mokken scale analysis

Refer to the supporting information available online for Mokken scale analysis.

2.6 | Person-item map

The ordering of the items along the latent trait (θ) was graphically displayed using a person-item map (PIM). The PIM shows the relationship between the estimated item location parameters and the estimated latent trait (ie, foot sensation), together with a histogram of the estimated latent trait values. The map provides useful graphical information on the ordering of items and the relationship between items and persons.

2.7 | Statistics

MSA was conducted using the R package Mokken (R Foundation for Statistical Computing). Other statistical analyses were carried out using IBM SPSS Statistics version 22.0 (IBM, Armonk, New York, USA). Missing data were replaced by imputed data using the procedure of expectation maximization, using 25 iterations. The Shapiro–Wilk test was used to assess normality. Because the majority of the variables significantly deviated from a normal distribution, we continued our analyses with nonparametric tests. Correlations between grading scale scores and demographic characteristics of the control population were investigated using Spearman coefficients. Using the Mann–Whitney U test, we compared differences between total item scores of RDF Study participants and controls without neuropathy as well as differences in grading scale scores between genders. Differences in grading scale scores of subjects with a second and third follow-up were assessed using the Friedman test. Spearman coefficients were used to determine the direction and magnitude of the correlations or differences between these change scores (as a measure of responsiveness), with the null hypothesis that no differences exist in foot sensation during follow-up. P < .05 (two-sided) was considered statistically significant.

3 | RESULTS

3.1 | General characteristics

A total of 416 diabetic patients with varying degrees of symptoms and loss of sensation were included in the RDF Study. Table S1 (online) shows the general characteristics of the patients. Fifty-two patients had a prior ulcer, and thirteen patients had a history of lower extremity amputation.

3.2 | Mokken scale analysis

Under the monotone homogeneity model, the automated item-selection procedure selected 39 of the original 42 items (Table S2 online). The items “Tinel sign left” and “Right” and “MNSI-positive symptoms” were not selected, so the 40-point scale ranges from 0 (no aberrant tests) to 39 (all tests aberrant). Scalability coefficient (H) values ranged from 0.354 to 0.713 (Table S2 online, third column), with a coefficient of scalability of H = 0.538, which indicates a strong scale—except for the item “Amputation left,” for which H > 0.8. The Molenaar-Sijtsma statistic, ρ = 0.964, suggested that the 39-item scale (RDF-39) is highly reliable. Of the 39 items, no items showed a violation of monotonicity. The most frequent aberrant items (eg, on S2PD and M2PD) represent early stages of sensory loss (Table S2 online, second column). Sensory functions were lost symmetrically, with items representing more distal test sites (eg, vibration sense at the interphalangeal joint) becoming aberrant before the proximal ones (eg, vibration sense at the medial malleolus).

Under the double monotonicity model, the manifest invariant item-ordering procedure selected 8 items violating IIO. The remaining scale consisted of 31 items (RDF-31) and had an HT coefficient of 0.581. H values ranged from 0.431 to 0.836 and featured an H = 0.550, which indicates a strong scale. The reliability statistic (ρ) was 0.958.

3.3 | Person-item map

A PIM showed that the 39 items of the RDF-39 could be classified into five core clusters (Figure S2 online). The S2PD cluster contained the items that were first becoming aberrant during the natural history of sensory loss, followed by a cluster of all M2PD items, vibration sense items, S1PD items, and items on prior ulceration/lower extremity amputation.
3.4 | Clinically applicable screening scale

Table S2 (online, sixth column) shows the item selection that we used to construct a clinically applicable screening scale, based on the 31-item scale. This 13-item scale (RDF-13) examines both extremities; items feature scalability coefficients ranging from 0.404 to 0.736 with $H = 0.551$, which indicates a strong scale. The reliability coefficient suggests that the scale is also reliable. Strong positive correlations have been found between the 39-item scale and the 31-item scale, $r_s = 0.993$, $P < 0.001$, and 31-item scale and 13-item scale, $r_s = 0.966$, $P < 0.001$. The item-response functions for the 13-item scale were plotted (Figure 1), showing their different discriminatory values along the sum score. As the latent trait increases (indicative of more severe sensory loss), so does the chance of obtaining an aberrant item test result. A clinically applicable scoring sheet for the respective scales is shown in Table S3 (online).

3.5 | Comparison to healthy controls

A total of 196 healthy volunteers, with a median age of 50.5 (interquartile range [IQR], 36.5-65.7) years, served as the control group—66 men (median age, 50.6 years; IQR, 37.6-64.1 years) and 130 women (median age, 50.0 years; IQR, 33.3-66.9 years). The median height for this group was 172.0 (IQR, 166.3-178.8) cm and the median weight 72.0 (IQR, 63.0-82.0) kg. Diabetic subjects were significantly older and heavier than controls ($P < 0.0005$; Table S1 online).

Figure 2 shows the 39-item sum-score distribution of diabetic RDF Study participants compared with the controls. Median total RDF-39 scores differed significantly between individuals in the control group (5.5; IQR, 3.0-10.8) and RDF Study subjects with diabetes (17; IQR, 9.0-22.0) ($P < 0.0001$).

Because diabetes-related lower extremity complications (ie, neuropathy symptoms, sensory loss, ulceration, and amputations) are the main drivers of RDF-39 scores in diabetic subjects, correlations between RDF-39 scores and demographic variables, such as age, height, and weight, were only explored in the control population. Significant positive correlations were found between RDF-39 scores and...
age ($r_s = 0.405, P < .0005$) and RDF-39 scores and weight ($r_s = 0.212, P = .003$). A nonsignificant positive trend was observed between RDF-39 scores and height, $r_s = 0.135, p = 0.059$. Nonsignificant differences in median RDF-39 scores (IQR) were found between males (6.0; IQR, 3.0–12.5) and females (5.0; IQR, 2.8–10.0) ($P = .563$).

### 3.6 Responsiveness of RDF-39

We conducted a Friedman test to determine whether there were differences in within-subjects’ RDF-39 scores, which were collected during follow-up of the RDF study ($n = 135$). RDF-39 scores at baseline (median, 17; IQR, 17–22), at the 1-year follow-up (median, 16; IQR, 16–24), and at 2-year follow-up (median, 18; IQR, 18–22) did not differ significantly ($\chi^2(2) = 1.536, P = .464$). There was a strong positive correlation between baseline and first follow-up RDF-39 scores ($r_s = 0.698, P < .0005$) and between first and second follow-up scores ($r_s = 0.697, P < .0005$).

### 4 DISCUSSION

This quantitative assessment of the categorical loss of pedal sensation of patients with diabetes has shown that the ability to sense S2PD, M2PD, vibration, and S1PD disappears in this order. This study has emphasized the added value of testing static and moving 2-point discrimination and highlights the importance of test locations in the screening of diabetic patients. Furthermore, the instruments (RDF-39, −31, and −13) captured the functional loss that is dictated by the pathophysiology of neuropathy.16,18,19

At present, no data are available on how a tuning fork, monofilament testing, and test locations compare or how these should be interpreted.11 By taking the site of screening into account, we observed that the first dorsal web and the lateral foot are the last of all sensibility tests and locations to become the least sensitive to the 10-g monofilament. This may have predictive value for future lower extremity complications because it suggests substantial deafferentation. Originally, the monofilament was studied as a prognostic indicator of ulceration and amputation.22,23 Nowadays, the validity of the monofilament examination to identify the presence of DSP is generally accepted.44,45 Because the onset of sensory loss is insidious, its diagnosis may be difficult. Several scoring systems for signs and symptoms of DSP have been developed, but they can be complex and time-consuming.46 Electrodiagnostic techniques do not assess all nerve fibers undergoing changes in diabetes and may be technically challenging on the plantar surface of the foot.47,48 It has been recommended that the diagnosis of DSP requires a test battery, with high sensitivity, that can detect early or mild forms in low-risk populations.49 Furthermore, it is important for population studies to possess screening tools that are reproducible, sensitive, and fast to carry out. The simple-to-use instruments used in the presented scales fulfill these criteria and are already being applied in clinical practice. These scales can quickly and reliably estimate skin sensation and may be easily implemented by nurses, nurse practitioners, and physicians treating diabetic patients.

MSA is also a flexible scaling method. More restrictive scaling methods, such as Rasch analysis, typically fit the data worse (ie, the data are not well-described by the Rasch model) and include relatively few items in the scale. However, the scales allow interval measurement.6 We believe ordinal measurement is sufficient for our study because the 40 ordinal levels of pedal sensory function produced by the 39-item scales are very informative. Most of the included items are indicators of large-fiber function; however, some items assessed small-fiber function (items “cold perception left” and “right”). We included these items to investigate how they become aberrant in the natural course of sensory loss, as compared with large-fiber function. Our data show that the ability to detect a cold stimulus decreases at a late stage of sensory loss, just before abnormal monofilament tests on the first dorsal web and lateral foot. However, the exact temporal sequence in which the different nerve fibers lose their functions is not fully understood.20,50 MSA may aid in this debate in future studies.

The results of our study confirm that a patient most often first loses sensation in the distal extremity, with vibration sense lost at the interphalangeal joint of the hallux before the medial malleolus. The scales also show that a patient loses sensation in both legs symmetrically, which is in line with the definition of DSP being a distal, symmetrical neuropathy.29 The item on numbness of the feet is positioned after the items on S2PD, M2PD, and vibration sense, which is of interest and suggests that patients seem unaware of the loss of these sensory modalities. At the same time, axonal density decreases, indicating that the feet are likely already at risk.51 The 39-item and 31-item scales (RDF-39 and -31) were developed for research purposes, yet they can also be used for patient-level measurements. The short 13-item (RDF-13) scale may help with individualized medical decision-making and may serve as a complement to the current prediction models for lower extremity complications.52–55

An automated item-selecting procedure ruled out the items “Tinel left” and “Right” and “MNSI-positive symptoms,” meaning that they did not pass the marginal test for fitting the monotone homogeneity model. These items do not hold a robust position in the natural course of the disease, as clinicians will recognize from daily practice—patients who have had an ulcer and patients without aberrant large-fiber function (eg, intact S2PD) may still complain of painful neuropathic symptoms. As these results suggest, subjective positive symptoms do not necessarily correlate with the degree of sensory loss, which is in contrast to negative symptoms experienced, such as “numbness,” which does have a robust position on the scale.

In our study, 44.9% (95% CI, 40.1%-49.7%) of diabetic patients exhibited signs of tibial nerve compression, as indicated by a positive Tinel sign at the tarsal tunnel.25 However, we also found that this sign had an uncertain place in the natural course of sensory loss; the Tinel sign items were not selected by the models and therefore were not included in the scales. This exclusion may be explained by the pathophysiology behind this diagnostic tool—demyelination and axonal sprouting elicit a positive sign, but a negative sign is reported when those phenomena have not yet occurred or when the nerve has been irreversibly damaged.56 Therefore, sensitivity/specificity calculations are not appropriate because they only can be interpreted when the degree of nerve damage is known.47,48

The population of the RDF Study has a wide variation sensory loss, with and without symptoms of DSP, resulting in a low risk of
spectrum bias. Therefore, we believe that the external validity of our findings is likely to be high. By comparing the distribution of the sum score of diabetic subjects to that of healthy volunteers without known neuropathy, we confirmed our hypothesis that the instruments correctly categorize patients’ sensation in the feet. However, due to prior dichotomization of items, with the threshold set at 8 mm for items on static and moving 2-point discrimination, we noted some aberrancy in the first 18 items (S2PD and M2PD) assessed in the healthy controls. Decline in foot sensation due to age was confirmed in our analysis, with age being the most important determinant. Only nonsignificant differences were observed between genders, which is in line with previous reports. The observed association between weight and foot sensation is presumably confounded by (components of) the metabolic syndrome, as it relates to polyneuropathy, and should be investigated in future studies. The most important risk factors for DSP relate to the duration of diabetes, control of glucose levels, and the existence of cardiovascular risk factors. We retained the null hypothesis that no differences existed in foot sensation in this time-frame. A follow-up of this cohort will reveal which time-frame is applicable on the progressive steps on the scale, as well as the associated risk for (re-)ulceration or amputation, per sum score.

In conclusion, MSA has revealed new dimensions in the use of current screening instruments in this diverse diabetic population. Based on the presented scales, clinicians may better categorize patients’ loss of sensation in their feet. Therefore, an individualized approach with recommendations regarding intervention strategies and patient information may be feasible.

ACKNOWLEDGMENTS

The authors thank Erwin Birnie PhD, Katharina B. E. Jorgensen, and N. Ahmad Aziz MD, PhD, for critical review of the manuscript and for their useful comments.

CONFLICT OF INTEREST

The authors declare no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

W.D.R. collected data, designed the study, performed the data analysis, and wrote the manuscript. M.H.A collected data, performed the data analysis, and wrote the manuscript. M.C.C. contributed to the discussion and reviewed/editors the manuscript. J.W.v.N. edited the manuscript. L.A.v.d.A. contributed to the data analysis, discussion, and edited the manuscript. J.H.C designed the study, contributed to the data analysis, and edited the manuscript.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal’s position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Willem D. Rinkel https://orcid.org/0000-0001-8137-3076

REFERENCES


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.