CHAPTER 4

Reliability of the Semmes Weinstein Monofilaments to measure coetaneous sensibility in the feet of healthy subjects

Susan Collins
Pepijn Visscher
Henrica C. de Vet
Wouter W.A. Zuurmond
Roberto S.G.M. Perez

Disability and Rehabilitation 2010;32(24):2019-27

Abstract

Purpose: Determine the intrarater-reliability, interrater-reliability and normal reference scores of the Semmes Weinstein Monofilament test (SWM) of the feet of healthy subjects. Additionally, the stability of the SWM for prospective use was assessed by determining systematic changes in sensory thresholds.

Methods: Interrater-reliability was assessed on 5 locations of the plantar side of both feet using monofilaments 1.65, 2.36, 2.44, 2.83, 3.22, 3.61, 3.84, 4.08, 4.31, 5.56, 6.65 in 60 healthy subjects by 2 or 3 investigators (test day 1). Intrarater-reliability and systematic changes in sensory thresholds were assessed 3 weeks later (test day 22) by one investigator.

Results: Median interrater-reliability for the 5 test locations for both feet was poor to moderate. Median intrarater-reliability was good for the left foot and poor to moderate for the right foot. Significantly lower median sensory thresholds were found for the first SWM measurement at test day 22 compared to the first and second measurement of test day 1. Given the observed reliability of the SWM, a normal sensory score for the feet was situated between monofilament 3.22 and 4.08.

Conclusions: The SWM are reliable when measured by one researcher. Systematic changes in sensory thresholds were observed, therefore, the stability of the SWM for use in prospective studies could not be verified.

Introduction

Illnesses affecting the somatosensory nervous system (1), such as Complex Regional Pain Syndrome type 1 (CRPS 1) (2), herpes zoster (3) and diabetic neuropathy (4) are characterized by sensory abnormalities such as allodynia, hyperesthesia and hypoesthesia, affecting coetaneous sensibility.

Various methods exist for determining coetaneous sensibility, of which the Semmes Weinstein Monofilament test (SWM) (5;6) is one of the most commonly used. The SWM consists of a kit of 20 nylon monofilaments attached to a rod. Each monofilament is of a different thickness, representing a logarithmic force on a base 10, with a range of 1.65 (the lowest) up to 6.65 (the highest), needed to bend the monofilament. The SWM has been used to measure sensory abnormalities in CRPS 1 (7;8), diabetic peripheral neuropathy and other chronic pain syndromes (9-11).

The SWM can be used to assess sensory disturbances in bilateral comparisons between affected and non-affected body regions, and in direct comparisons with normal reference standards (12) as specified by the SWM manufacturer for the upper extremity (see appendix 1). Although clinimetric properties of the SWM have been established for the upper extremity (13-16), information with respect to the reliability of the SWM and normal reference is lacking for the lower extremity. Physical influences from functional use of the lower extremities, in particular for the plantar side of the feet, differ substantially from those exhibited to the upper extremity, and may therefore lead to a different quality in sensory perception.

The purpose of the present study is to assess the intrarater and the interrater-reliability, and to obtain a normal reference sensory score of the SWM for the feet of healthy individuals. Furthermore, in order to assess the stability of the SWM for use in prospective studies, systematic changes in sensory thresholds measured with the SWM were assessed.

Methods

This study evaluating the interrater-reliability of the SWM was preceded by 2 preparation studies, in which in the first preparation study the monofilaments required were selected. In the second preparation study, intrarater-reliability was assessed on the selected monofilaments. The results of the final interrater-reliability study will be described in more detail. For the 2 preparation studies and the final interrater-reliability study, the following test procedure and statistical analyses were performed.

The test procedure

The tests were performed in a quiet and distraction free room under environmentally stable conditions, and were performed by trained researchers according to a standardized protocol.

For acclimatization purposes, ten minutes before the start of the measurements, subjects were asked to take both shoes and socks off and to lie down on the treatment bench. In addition, a short questionnaire about the actual health status of the subjects was assessed and test instructions were given to the subjects.

Temperature of the feet was measured on 5 locations of the dorsum of the feet with an infrared thermometer (First Temp Genius®) and was averaged. To control for possible decrease in temperature of the feet, mean foot temperature measured before the first test was compared with mean foot temperature measured after the last test (17).

As there is no consensus about the used SWM test locations (18), 5 SWM testing areas of the plantar side of the foot, representing various sensibilities of the foot, were selected. Test locations were: the distal phalanx dig.1, the distal phalanx dig.2, the distal phalanx dig.5, the arcus plantaris medialis and the arcus plantaris lateralis. All monofilaments were applied three times consecutively, except for filaments 3.84 and higher, which were applied once according to the manufacturer's instructions. Subjects were instructed to close their eyes during the test. All monofilaments were applied 1.0-1.5 seconds to the test locations, starting with the smallest monofilament. Subjects were asked to name the exact location where a monofilament was detected. If a monofilament was perceived correctly on a test location, the filament representing a specific force was noted as the subject's sensory score. If a monofilament was not detected, the researcher tested the next monofilament in rank. Subjects were kept unaware of the measurement sequence, and were given no indication as to whether a monofilament was identified correctly.

<u>Preparation study 1. Selection of monofilaments and intra-rater assessment</u>

In the first phase of this study, the intrarater-reliability of the SWM was tested in the feet of 50 (20 male and 30 female) healthy subjects (age: mean 40.4, SD 11.2 years). Five locations of the plantar side of the feet were tested bilaterally. To obtain a reference value for future comparisons, the monofilaments 1.65, 2.83, 3.22, 3.61, 4.31, 4.56 and 6.65 were used, representing sensory ranges specified by the manufacturer (see appendix 1). Sensory thresholds were found to be situated between monofilament 3.61 and 4.31 (3.75; IQR 3.52-4.24 and 3.75; IQR 3.52-4.03, for the right and left foot respectively). Women had significantly lower median

sensory scores than men (women vs. men for the left foot: 3.66; IQR 3.38-4.03, 3.95; IQR 3.67-4.31, p= 0.038 and women vs. men for the right foot: 3.61; IQR 3.45-3.99, 3.89; IQR 3.61-4.31, p= 0.025). Intrarater-reliability was moderate for the right foot (ICC =.66; IQR.57-79) and very good for the left foot (ICC =.84; IQR.68-.91). Consequently, in order to obtain a more sensitive estimation of the sensory threshold between filaments 3.61 and 4.31, monofilaments 3.84, 4.08 and 4.17 were included.

Preparation study 2. Second testing of monofilaments and intrarater-reliability assessment

In the second preparation study, the intrarater-reliability of the augmented set of monofilaments was assessed. Ten male and 10 female subjects (mean age 40.4, SD 10.3 years) from the sample used in the first phase were retested after 4 weeks. The intrarater-reliability was found to be good for both feet (right foot ICC:0.83; IQR 0.73-0.86, left foot ICC:0.86; IQR 0.83-0.92). Contrary, to the findings in the first preparation study significantly lower sensory thresholds were found (3.41; IQR 3.22-3.58, 3.37; IQR 3.22-3.47, p<0.000 for the left and right foot, respectively), without significant differences between both sexes.

Final interrater-reliability study

Based on preparation study 1 and 2, the final study was started in which the interrater-reliability and systemic differences in sensory thresholds over time were assessed. The same filaments were used as in preparation study 2 with the addition of monofilaments 2.36 and 2.44. These monofilaments were added to the testing protocol, due to the finding of significantly lower sensory thresholds in preparation study 2 compared to preparation study 1. Monofilament 4.17 was removed from the testing protocol because it appeared to have no added value above included monofilaments. As different subjects and researchers participated in this interrater-reliability study compared to the preparation studies, the intrarater-reliability was determined again.

Subjects

Sixty healthy subjects were recruited from hospital staff and medical students. Inclusion criteria were: 1. Age between 20 to 70 years; 2. No impairments in the lower extremity; 3. No recent traumas or operations in the testing area; 4. Written informed consent. Exclusion criteria were: 1. Sensibility or pain complaints of the feet; 2. Skin abnormalities (e.g., scars) in the testing areas; 3. Diagnosed with

diabetes mellitus and/or polyneuropathy; 4. Lower back complaints in combination with radicular pain in the leg; 5. Other complaints that could influence sensibility of the feet. Subjects were divided in 5 age subgroups: 20-30 years, 31-40 years, 41-50 years, 51-60 years and 61-70 years. All subgroups contained 10 subjects except for subgroup 41-50 years which contained 20 subjects, as the prevalence of chronic pain disorders affecting the somatosensory system is higher in this age group. In the CRPS 1 population of our pain clinic a gender distribution of 2:3 is found. To mimic this gender distribution each subgroup contained 40% men and 60% women.

Interrater-reliability (test day 1)
In order to test the interrater-reliability of the SWM, measurements were repeated by a second and/or third researcher. The second and third test was performed 5 minutes after the preceding test. The testing order of the researchers was randomized.

Intrarater-reliability (test day 22)
Three weeks after test day 1 (test day 22), additional SWM measurements were performed by only 1 researcher. To obtain information about the intrarater-reliability of the SWM, measurements were repeated immediately after finishing the first measurement.

Systematic changes in sensory thresholds
Systematic changes in sensory thresholds over time were examined by comparing the sensory thresholds measured on test day 1 with the sensory thresholds measured on test day 22. As there were no interventions between the measurements at day 1 and 22, the sensory thresholds were expected to have about the same value.

Statistical analysis
The data were processed and analyzed with SPSS 15. Intrarater-reliability and interrater-reliability (for 2 and 3 researchers) were calculated for the 5 SWM test locations individually, using the Intraclass Correlation Coefficient (ICC). Subsequently, the 5 individual ICC's were used to calculate the median intrarater and interrater-reliability for the whole foot. The following cut-off-points for the ICC were used: <0.50 poor reliability, 0.50-0.75 moderate reliability, >0.75 good reliability (19;20). Median interrater and intrarater-reliability were determined for the whole sample, and for men and women. In order to control for the influence of temperature decrease on sensory perception, subgroup analyses were performed for subjects

with a decrease in skin temperature, using cut off point of 0.6°C for 2 raters, or 1.0°C for 3 raters. These cutoff points were used, because a skin temperature difference of respectively 0.6°C and 1.0°C between the affected and unaffected extremity in CRPS 1 patients was shown to be a predictor for the diagnosis of CRPS 1 (21). Data were presented as mean (± SD) or medians (IQR) when appropriate.

Wilcoxon signed rank test was used to compare ICC scores of subjects with low and high decrease of skin temperature and to compare sensory thresholds of test day 1 with those of test day 2. Friedman test was used to compare reliability scores between different test locations. P values < 0.05 were considered statistically significant.

The standard error of measurement (SEM= σ√1-ICC) was obtained from the intrarater-reliability and was used to calculate the minimal detectable change (i.e. the amount of change greater than measurement error) in sensory thresholds (22). Minimal detectable change in sensory thresholds was defined as the change between consecutive measurements exceeding the interval: 1.96 * √2 * SEM. These 1.96 * √2 * SEM corresponds to the limits of agreement in a Bland and Altman method (23).

Results

Sixty healthy subjects were included with a mean age of 44.0 (SD 13.3).

Interrater-reliability

Thirty-four subjects were tested by 2 researchers and 26 by 3 researchers. Table 1 shows the median interrater-reliability calculated between the researchers (3 researchers, the first and second researcher, the first and third researcher and the second and third researcher, respectively) for the right and left foot. Median interrater-reliability for both the right and left foot was poor to moderate for the whole sample and for women and men separately.

Table 1: Median interrater-reliability for the three raters and per pair

Researchers		1, 2 and 3		1 and 2		1 and 3		2 and 3	
		N	ICC (IQR)	N	ICC (IQR)	N	ICC (IQR)	N	ICC (IQR)
All	**R**	26	0.44 (0.40-0.47)	59	0.49 (0.39-0.59)	26	0.36 (0.31-0.53)	26	0.50 (0.31-0.72)
	L	26	0.41 (0.16-0.61)	59	0.42 (0.39-0.56)	26	0.36 (-0.09-0.56)	26	0.52 (0.29-0.72)
Women	**R**	15	0.33 (0.17-0.52)	35	0.58 (0.32-0.65)	15	0.38 (0.23-0.53)	15	0.27 (0.09-0.56)
	L	15	0.18 (0.10-0.42)	35	0.41 (0.32-0.58)	15	0.00 (-0.15-0.35)	15	0.31 (0.11-60.1)
Men	**R**	11	0.56 (0.30-0.65)	24	0.42 (-0.21-0.54)	11	0.28 (-0.40-0.72)	11	0.58 (0.44-0.90)
	L	11	0.46 (0.30-0.47)	24	0.45 (0.41-0.63)	11	0.40 (0.12-0.71)	11	0.48 (0.34-0.83)

R: right foot, L: left foot. Median taken of 5 test locations, IQR: inter quartile range.

Poor to moderate ICC scores were found for the separate test locations (table 2). The 5 test locations measures in the left foot differed significantly in reliability scores, with dig. 2 and 5 showing the highest reliability and the arcus plantaris medialis the lowest (Friedman test, p = 0.015). No significant difference was found between the test locations for the right foot (Friedman test, p = 0.699).

Table 2: ICCs for interrater-reliability per location for the whole sample

		Dig. 1	Dig. 2	Dig. 5	Medialis	Lateralis
1, 2 and 3	**R**	0.40	0.45	0.44	0.40	0.48
	L	0.41	0.54	0.68	0.07	0.25
1 and 2	**R**	0.45	0.33	0.49	0.63	0.54
	L	0.40	0.42	0.63	0.37	0.48
1 and 3	**R**	0.16	0.36	0.25	0.38	0.68
	L	0.46	0.37	0.61	-0.19	0.01
2 and 3	**R**	0.74	0.50	0.69	0.24	0.38
	L	0.30	0.81	0.68	0.27	0.53

R: right foot, L: left foot.

The influence of decrease in skin temperature on the reliability of the SWM is shown in table 3 (N=57, missing n=2). Most of the subjects tested by 3 researchers had a skin temperature decrease of more than 0.6°C. No significant difference in reliability scores was found between the subjects with skin temperature decrease higher or lower than 0.6°C (p = 0.69) or 1.0°C (p = 0.50), respectively.

Table 3: Influence of temperature on reliability on median ICC

2 researchers		**N**	**ICC (IQR)**
< 0.6°	R	10	0.55 (0.50-0.80)
> 0.6°	R	23	0.51 (0.19-0.60)
< 0.6°	L	6	0.40 (0.17-0.64)
> 0.6°	L	27	0.41 (0.34-0.62)
3 researchers		**N**	**ICC (IQR)**
< 1.0°	R	5	0.60 (0.08-0.80)
> 1.0°	R	19	0.33 (0.26-0.47)
< 1.0°	L	6	0.42 (-0.36-0.89)
> 1.0°	L	18	0.07 (-0.05-0.22)

R: right foot, L: left foot. Median taken of 5 test locations, IQR: inter quartile range.

Intrarater- reliability

Reliability scores for the whole sample and individual test locations are described in table 4. Good intrarater-reliability scores were found for the left foot for the whole sample, women and men separately, while poor to moderate reliability scores were found for the right foot.

Reliability scores were significantly higher in the left foot compared to the right foot for the whole sample ($p = 0.043$) and women ($p = 0.043$) and tended to be higher in men ($p = 0.08$). No significant differences in intra-rater-reliability were found between different test locations for the right (Friedman test, $p = 0.569$) and left foot (Friedman test, $p= 0.549$).

Bland and Altman plots (figures 1 and 2) for the left and right foot depict abovementioned differences. Negative mean differences between the first and second measurements indicate slightly higher sensory thresholds during the second measurement. No relationship between the magnitude of measurements and size of the difference could be observed.

Table 4: median intrarater-reliability for the 5 test locations and ICC per location

			Median ICC	**ICC per location**				
		N	**ICC (IQR)**	**Dig. 1**	**Dig. 2**	**Dig. 5**	**Medialis**	**Lateralis**
All	R	57	0.52 (0.42-0.58)	0.52	0.36	0.56	0.61	0.48
	L	57	0.78 (0.68-0.83) *	0.80	0.78	0.84	0.78	0.69
Women	R	34	0.56 (0.53-0.67)	0.72	0.63	0.56	0.52	0.54
	L	34	0.90 (0.80-0.94) *	0.86	0.91	0.90	0.55	0.96
Men	R	23	0.41 (0.23-0.65)	0.37	0.10	0.58	0.73	0.41
	L	23	0.83 (0.48-0.86)	0.83	0.63	0.87	0.85	0.32

R: right foot, L: left foot. IQR: inter quartile range. * Significantly higher reliability scores in the left foot compared to the right foot ($p< 0.05$)

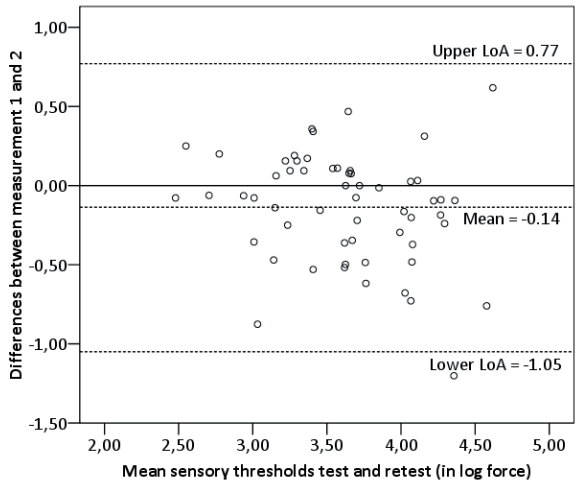

Figure 1: Bland and Altman plot for the differences between measurement 1 and 2 against mean test and retest sensory thresholds for the right foot

LoA: limits of agreement

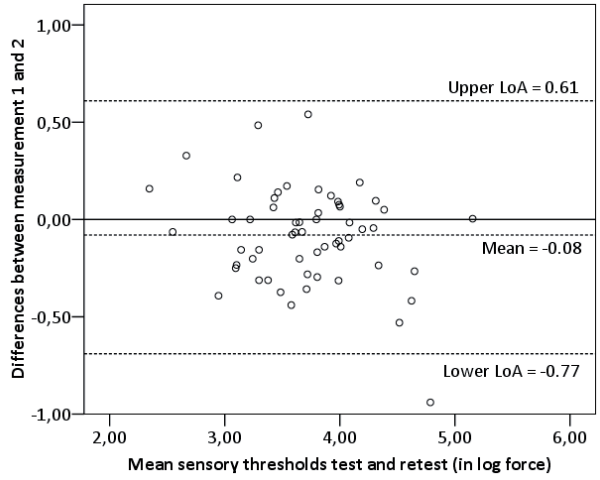

Figure 2: Bland and Altman plot for the differences between measurement 1 and 2 against mean test retest sensory thresholds for the left foot

LoA: limits of agreement

Systematic changes in sensory scores

Sensory thresholds measured during test day 1 and test day 22 for the whole sample, men and women separately are presented in table 5. For the whole sample, sensory thresholds were found to lie between monofilaments 3.22 and 4.08, without statistically significant differences between men and women.

Statistically significant lower sensory thresholds were found for the first SWM measurement at test day 22 compared to the first and second measurement of test day 1. No statistically significant differences were found in sensory thresholds

between measurement 2 on test day 22 compared to measurement 1 and 2 of test day 1.

Table 5: Median sensory thresholds for the first, second and third measurement of test day 1 and for the first and second measurement of day 22 for the whole sample, women and men separately

		Test day 1			Test day 22	
Measurement		**1**	**2**	**3**	**1 (test)**	**2 (retest)**
All	R	3.69 (3.37-4.04)*	3.75 (3.38-4.12)*	3.63 (3.34-3.86)	3.59 (3.30-3.86)	3.67 (3.23-4.12)
	L	3.74 (3.22-4.07)	3.82 (3.47-4.11)*	3.61 (3.26-3.91)	3.64 (3.25-4.03)	3.80 (3.38-4.08)
Women	R	3.64 (3.31-4.04)*	3.69 (3.30-4.13)*	3.57 (3.22-3.80)	3.59 (3.30-3.84)	3.63 (3.23-4.03)
	L	3.69 (3.22-4.08)	3.81 (3.39-4.03)	3.39 (3.22-3.83)	3.64 (3.20-4.00)	3.72 (3.22-4.03)
Men	R	3.76 (3.54-4.17)*	3.81 (3.48-4.22)*	3.70 (3.45-3.98)	3.58 (3.30-3.89)	3.73 (3.22-4.27)
	L	3.83 (3.30-4.07)	3.89 (3.58-4.22)	3.77 (3.45-3.94)	3.80 (3.45-4.17)	3.86 (3.39-4.22)

Sensory thresholds in median (IQR). *Statistically significant higher sensory thresholds at day 1 compared to measurement 1 on day 22 ($p < 0.05$).

The SEM values and limits of agreements are presented in table 6. Values are slightly higher for males compared to females and right foot compared to the left. For the total sample, the minimal detectable change in sensory threshold (i.e. changes exceeding the interval: 1.96 * √2 * SEM) does not exceed 0.91 points on the log force scale (see Appendix A). This means that, given the intarater-reliability found in this study, a change less than 0.91 in a person is probably due to measurement error, but changes larger than 0.91 points might be considered real changes.

Table 6: Maximum sensory thresholds, mean sensory thresholds, ICC´s, SEM and limits of agreement for the whole sample, women and men separately

	Right foot					Left foot				
	Max sensory Threshold †	Sensory Thresholds ‡	ICC §	SEM ‖	LoA ¶	Max sensory Threshold †	Sensory Thresholds ‡	ICC §	SEM ‖	LoA ¶
All	4.93	3.56 (0.49)	0.52	0.33	± 0.91	5.16	3.68 (0.53)	0.78	0.25	± 0.69
women	4.32	3.55 (0.46)	0.56	0.31	± 0.85	4.52	3.59 (0.51)	0.91	0.15	± 0.42
men	4.93	3.56 (0.55)	0.41	0.42	± 1.16	5.16	3.80 (0.54)	0.83	0.22	± 0.61

†Highest sensory score for the whole feet. ‡Mean sensory thresholds (SD). §ICC: Intrarater-reliability.
‖SEM: standard error of measurement.
¶LoA: Limits of agreement defined as: 1.96 *√2 *SEM.

Discussion

Poor to moderate interrater-reliability for the SWM tested for the feet was found in the present study. Our results are in line with those from the study by Rozental et al. (24), who assessed interrater-reliability of the SWM measured in hands of

healthy volunteers and found low reliability scores. However, in contrast to our and Rozental et al's findings, other studies evaluating the interrater-reliability of the SWM (11;15;16;25-27) found much higher reliability scores. These studies (11;15;16;25-27), however, have evaluated the interrater-reliability of the SWM in patients. Reliability parameters (such as the ICC) are highly dependent on the heterogeneity of a study sample (22). Therefore, the lower interrater-reliability thresholds found in our and Rozental et al's (24) study are, due to the less inter individual variation of sensory thresholds in healthy subjects compared to patients, to be expected.

Although the researchers in our study were all well trained and the SWM test was performed according to a strict and standardized assessment protocol, minimal differences in applied force between assessors may have influenced results (28). Examining median sensory thresholds for the 3 examiners (SC, PV and RP) revealed that slightly higher (non significant) sensory thresholds were obtained by SC compared to PV and RP (data not presented). Exclusion of the measurements of SC from the analyses did not significantly improve reliability scores. An alternative explanation may lie in actual differences in perception of sensory stimulation by patients in consecutive measurement.

We are of opinion that testing 2 extremities represents best the current clinical practice in which the non-affected extremity is compared to the affected extremity. Whereas most studies evaluating the interrater-reliability of the SWM evaluated only 1 extremity, in the present study we assessed both feet, resulting in a longer test duration. Prolonged testing duration may have influenced subject concentration, leading to reduced alertness for sensory stimulation. Furthermore, the increase in testing duration appears to have led to decreased temperature of the feet due to reduced circulation requirements. Cooling down of the feet may lead to temporary numbed coetaneous sensory fibers resulting in lower sensitivity scores (29). Although in our sample lower sensitivity scores were not observed as the feet cooled down (data not shown), we do think that studies using the SWM should take the temperature of the measured extremities into account.

Differences in the used filaments may also contribute to the differences in interrater-reliability. Most studies used filaments that represented the upper limit of each sensory range (2.83: normal touch, 3.61: diminished light touch, 4.31: diminished protective, 4.56: loss protective and 6.65: deep pressure only) as stated by the manufacturer. We have added the monofilaments 1.65, 2.36, 2.44, 3.22, 3.84 and 4.08 to our testing protocol, in order to increase sensitivity of the test. However, detection of individual monofilaments may have been hampered due to smaller

inter-monofilament differences in relation to variable discriminative properties of sensory receptors.

Furthermore, differences in the tested locations may have influenced overall reliability scores. On the plantar side of the feet interrater-reliability scores were in general highest in digi 2 and digi 5 compared to the other test locations. Although these difference were not statistically significant, this may indicate that some test sites are more reliable than others. Further evaluation of alternative test sites (e.g., the dorsum site of the foot) may therefore provide more reliable testing areas for the feet.

In contrast to the interrater-reliability, the intrarater-reliability was found to be moderate to good for the whole sample, which appeared to be consistent across the preparation studies as well as the final study. Remarkable were the significantly higher intrarater-reliability scores for the left foot compared to the right foot. In preparation study 1, we also found a slightly lower reliability score for the right foot. In both cases, the right foot was tested first in line according to protocol. Possibly, subjects were more reserved in response due to the unfamiliarity with the perception of the filaments.

Furthermore, lower intrarater-reliability scores were found in the final study compared to those found in preparation study 1, despite equal testing circumstances. A possible explanation for the differences between the preparation studies and the final study is that different subjects and different researchers participated in both studies. These findings underline our observations that inter-individual variation on subject and observer level contribute to reliability of SWM measurements for the feet.

In this sample of subjects, sensory thresholds for the whole foot were found to lie between monofilaments 3.22 and 4.08. In a study performed on healthy subjects by Kets et al. (30), monofilament 4.31 was found to be the first filament perceived on all locations of the foot. A mean sensory threshold score of 3.63 was found in the foot of healthy subjects by Jeng et al. (31). Furthermore, mean sensory thresholds corresponding with monofilament 3.61 were found in the feet of healthy subjects by Kemler et al. (12). Based on the latter and the concurrence in median sensory threshold scores found in the final study and first preparation study, we suggest that normal sensitivity of the feet lies between monofilaments 3.22 and 4.08.

Systematic changes in sensory thresholds were observed during test day 1 and 22, which we could not explain. Whether these differences are related to the subjects or the SWM instrument needs to be established more thoroughly. Therefore, for now, the stability of the SWM assessments of the feet for use in prospective studies cannot be verified.

Conclusion

This study found limited interrater-reliability of the SWM, suggesting that SWM thresholds scores obtained by different researchers are less comparable. According to our present data, we conclude that SWM measurements of the plantar side of the feet are reliable when measured by one researcher. In addition, in this sample of subjects, a normal sensory score is situated between monofilament 3.22 and 4.08. Whether this score is generally applicable as a normal reference value needs to be further elucidated.

The SWM are capable to detect true changes in sensory thresholds, however, is not clear whether the SWM is stable enough to be used in prospective studies. Therefore, the stability of the SWM needs to be further evaluated. In addition, future studies should also consider alternative test locations of the feet (e.g., dorsum side) or lower extremity for the SWM.

Considerations for clinical use of the SWM for the feet

The SWM is an inexpensive, noninvasive, easy to use instrument that quantifies pressure detection thresholds of the skin and is therefore of value to the clinical setting. The clinical value of the SWM when used in the feet, however, can be (negatively) influenced by some aspects, which those working in the field should consider.

Sensory disturbances can be assessed with the SWM by comparing affected with non-affected body regions, or by direct comparison with normal reference standards. Direct side to side comparison is preferred (12), but in cases where comparison within the same subject is not possible, age and sex categorized normal sensory thresholds for the plantar side of the foot as determined in this study can be used.

No agreement exist about the number of monofilaments and which filaments should be used in SWM testing of the feet. In case of within subject comparison, we recommend to use a broad range of monofilaments. However, the use of all 20 monofilaments is not recommended as it is time consuming, and my lead to possible loss of subjects' concentration and cooling down of the feet. For comparison with normal reference values, we suggest to test monofilaments in the middle of the normal reference range, and those outside this range. In this case of our study, for use in the plantar side of the feet, monofilaments 3.22 to 4.08, 1.65 to 2.83 and 4.17 to 6.65 would be appropriate to represent respectively, normal, hyperesthesic and hypoesthesic sensory values.

As there is no consensus regarding specific test localizations, we recommend to

choose test areas representative for the disorder. In disorders in which specific peripheral nerves are affected, locations representing these nerves can be chosen. In disorders without evident nerve damage, test locations representing different levels of sensibility (for instance both dorsum and plantar side of the feet) can be selected.

Normal longitudinal variations in sensory thresholds for the plantar side of the feet are evident, and have to be taken into consideration when assessing changes in the course of time**.** Furthermore, SWM measurements in the feet performed by different researchers cannot be compared with each other, due to limited interrater-reliability. The SWM measurements can be reliably used for diagnosing sensory abnormalities (such as hyperesthesia and hypoesthesia) of the plantar side feet when assessed by one researcher.

Appendix

Sensory ranges for the hand as specified by Semmes Weinstein pressure Aesthesiometer manufactured by Smith & Nephew (USA)

Log force	Grams	Sensory range
1.65	0.0045	Normal
2.36	0.0230	
2.44	0.0275	
2.83	0.0677	
3.22	0.1660	Diminished light touch
3.61	0.4082	
3.84	0.6958	Diminished protective sensation
4.08	1.1940	
4.17	1.4940	
4.31	2.0520	
4.56	3.6320	Loss of protective sensation
4.74	5.500	
4.93	8.650	
5.07	11.70	
5.18	15.00	
5.56	29.00	
5.88	75.00	
6.10	127.0	
6.45	281.5	
6.65	447.0	

Reference List

(1) Dworkin RH, Backonja M, Rowbotham MC, Allen RR, Argoff CR, Bennett GJ, et al. Advances in neuropathic pain - Diagnosis, mechanisms, and treatment recommendations. Archives of Neurology 2003 Nov;60(11):1524-34.

(2) Birklein F, Riedl B, Sieweke N, Weber M, Neundorfer B. Neurological findings in complex regional pain syndromes - analysis of 145 cases. Acta Neurologica Scandinavica 2000 Apr;101(4):262-9.

(3) Kost RG, Straus SE. Postherpetic Neuralgia -- Pathogenesis, Treatment, and Prevention. N Engl J Med 1996 Jul 4;335(1):32-42.

(4) Dyck PJ, Dyck PJ, Larson TS, O'Brien PC, Velosa JA. Patterns of quantitative sensation testing of hypoesthesia and hyperalgesia are predictive of diabetic polyneuropathy: a study of three cohorts. Nerve growth factor study group. Diabetes Care 2000 Apr 1;23(4):510-7.

(5) Bell-Krotoski JA, Fess EE, Figarola JH, Hiltz D. Threshold detection and Semmes-Weinstein monofilaments. J Hand Ther 1995 Apr;8(2):155-62.

(6) Bell-Krotoski J. Advances in sensibility evaluation. Hand Clin 1991 Aug;7(3):527-46.

(7) Rommel O, Malin JP, Zenz M, Janig W. Quantitative sensory testing, neurophysiological and psychological examination in patients with complex regional pain syndrome and hemisensory deficits. Pain 2001 Sep;93(3):279-93.

(8) Sethna NF, Meier PM, Zurakowski D, Berde CB. Cutaneous sensory abnormalities in children and adolescents with complex regional pain syndromes. Pain 2007 Sep;131(1-2):153-61.

(9) Cruccu G, Anand P, Attal N, Garcia-Larrea L, Haanpaa M, Jorum E, et al. EFNS guidelines on neuropathic pain assessment. European Journal of Neurology 2004 Mar;11(3):153-62.

(10) Kamei N, Yamane K, Nakanishi S, Yamashita Y, Tamura T, Ohshita K, et al. Effectiveness of Semmes-Weinstein monofilament examination for diabetic peripheral neuropathy screening. Journal of Diabetes and its Complications 2005;19(1):47-53.

(11) Valk GD, de Sonnaville JJJ, van Houtum WH, Heine RJ, van Eijk JTM, Bouter LM, et al. The assessment of diabetic polyneuropathy in daily clinical practice: Reproducibility and validity of Semmes Weinstein monofilaments examination and clinical neurological examination. Muscle & Nerve 1997 Jan;20(1):116-8.

(12) Kemler MA, Schouten HJA, Gracely RH. Diagnosing sensory abnormalities with either normal values or values from contralateral skin - Comparison of two approaches in complex regional pain syndrome I. Anesthesiology 2000 Sep;93(3):718-27.

(13) Jerosch-Herold C. Assessment of sensibility after nerve injury and repair: A systematic review of evidence for validity, reliability and responsiveness of tests. Journal of Hand Surgery-British and European Volume 2005 Jun;30B(3):252-64.

(14) Marx RG, Hudak PL, Bombardier C, Graham B, Goldsmith C, Wright JG. The reliability of physical examination for carpal tunnel syndrome. The Journal of Hand Surgery: Journal of the British Society for Surgery of the Hand 1998 Aug;23(4):499-502.

(15) Novak CB, Mackinnon SE, Williams JI, Kelly L. Establishment of Reliability in the Evaluation of Hand Sensibility. Plastic and Reconstructive Surgery 1993 Aug;92(2):311-22.
(16) van Brakel WH, Khawas IB, Gurung KS, Kets CM, van Leerdam ME, Drever W. Intra- and inter-tester reliability of sensibility testing in leprosy. Int J Lepr Other Mycobact Dis 1996 Sep;64(3):287-98.
(17) Oerlemans HM, Perez RSGM, Oostendorp RAB, Goris RJA. Objective and subjective assessments of temperature differences between the hands in reflex sympathetic dystrophy. Clinical Rehabilitation 1999 Oct;13(5):430-8.
(18) Lee SY, Kim HH, Choi SH, Park YS, Kim YJ, Cho BM. Clinical usefulness of the two-site Semmes-Weinstein monofilament test for detecting diabetic peripheral neuropathy. Journal of Korean Medical Science 2003 Feb;18(1):103-7.
(19) Altman DG. Practical statistics for medical research. London: Champman and Hall; 1991.
(20) Portney LG, Watkins MP. Foundations of clinical practice: applications to practice. Norwalk: Appleton and Lange; 1993.
(21) Bruehl S, Lubenow TR, Nath H, Ivankovich O. Validation of thermography in the diagnosis of reflex sympathetic dystrophy. Clinical Journal of Pain 1996 Dec;12(4):316-25.
(22) De Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. Journal of Clinical Epidemiology 2006 Oct;59(10):1033-9.
(23) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-10.
(24) Rozental TD, Beredjiklian PK, Guyette TM, Weiland AJ. Intra- and interobserver reliability of sensibility testing in asymptomatic individuals. Annals of Plastic Surgery 2000 Jun;44(6):605-9.
(25) Anderson AM, Croft RP. Reliability of Semmes Weinstein monofilament and ballpoint sensory testing, and voluntary muscle testing in Bangladesh. Lepr Rev 1999 Sep;70(3):305-13.
(26) Birke JA, Brandsma JW, Schreuders TAR, Piefer A. Sensory testing with monofilaments in Hansen's disease and normal control subjects. International Journal of Leprosy and Other Mycobacterial Diseases 2000 Sep;68(3):291-8.
(27) Herron ML, Kar S, Beard D, Binfield P. Sensory dysfunction in the great toe in hallux valgus. J Bone Joint Surg Br 2004 Jan 1;86-B(1):54-7.
(28) Bell-Krotoski J, Tomancik E. The repeatability of testing with Semmes-Weinstein monofilaments. J Hand Surg [Am] 1987 Jan;12(1):155-61.
(29) Doeland HJ, Nauta JJ, van Zandbergen JB, van der Eerden HA, van Diemen NG, Bertelsmann FW, et al. The relationship of cold and warmth cutaneous sensation to age and gender. Muscle Nerve 1989 Sep;12(9):712-5.
(30) Kets CM, van Leerdam ME, van Brakel WH, Deville W, Bertelsmann FW. Reference values for touch sensibility thresholds in healthy Nepalese volunteers. Lepr Rev 1996 Mar;67(1):28-38.

(31) Jeng C, Michelson J, Mizel M. Sensory thresholds of normal human feet. Foot Ankle Int 2000 Jun;21(6):501-4.