AN ADAPTIVE PROCEDURE FOR ROUTINE MEASUREMENT OF LIGHT-TOUCH SENSITIVITY THRESHOLD

ANNE D. BERQUIN, MD, PhD,[1] VALÉRIA LIJESEVIC, PT,[1] SERGE BLOND, MD,[2] and LÉON PLAGHKI, MD, PhD[1]

[1] Physical Medicine and Rehabilitation, Cliniques Universitaires St. Luc, Avenue Hippocrate 10/1650, 1200 Brussels, Belgium
[2] Neurosurgery Department, Centre Hospitalier Régional Universitaire de Lille, Lille, France

Accepted 2 February 2010

ABSTRACT: Quantitative sensory testing with Semmes–Weinstein monofilaments suffers from several pitfalls. Our aims were to assess the reliability of these filaments for touch–pressure threshold detection, develop and validate a rapid and accurate procedure for measurements at the bedside, and establish normative data. After calibration of the monofilaments, an adaptive staircase algorithm was validated and used to establish normative data in healthy subjects. Calibration showed significant differences between manufacturer- and investigator-produced data. The relative humidity significantly affected the force exerted by the filaments. The adaptive procedure showed good accuracy and substantial time-saving. Touch–pressure thresholds showed significant gender differences (mean ± 2 SD for females/males: 2.82–12.3/3.09–17.78 g/mm^2). The influence of body site and age is small. Accurate use of Semmes–Weinstein monofilaments requires prior calibration, correction for humidity, and use of a validated procedure. In this study we provide normative data that can be used with our algorithm.

Muscle Nerve **42:** 328–338, 2010

In patients with neurological conditions, a comprehensive assessment of sensory abnormalities can be very useful. In diabetic polyneuropathy, for example, the occurrence of foot ulcerations is correlated with impaired pressure sensation.[1] In neuropathic pain, different patterns of sensory changes may reflect distinct pathophysiological mechanisms, which could justify a differentiated therapeutic approach.[2] Quantitative sensory testing (QST) permits the detection of both hyper- and hypophenomena (i.e., allodynia as well as hypesthesia[3]) and allows quantification of sensory changes, thus complementing electrodiagnostic methods.[1,4,5]

QST suffers from several pitfalls. Testing algorithms and devices are heterogeneous, the results critically depend on subject cooperation, and normative data are scarce and highly influenced by methodological issues. Therefore, the need for standardized clinically applicable methods and normative data has been highlighted by several investigators.[1,5]

Measurement of touch–pressure detection thresholds usually involves the use of Semmes–Weinstein monofilaments, derived from von Frey hairs, despite several caveats regarding their use. The stimulus waveform is variable, extraneous movements may occur,[6] and calibration experiments show large discrepancies between the force values indicated by the manufacturer and the actual data.[7] Moreover, variations in relative humidity significantly affect the applied force.[8] In spite of these problems, the use of Semmes–Weinstein monofilaments is supported by the Peripheral Neuropathy Association,[6] probably because of the simplicity of their use. However, it is essential to perform careful studies that include calibration experiments. The relative humidity must also be taken into account. We have found five studies that specifically assess tactile sensory thresholds with Semmes–Weinstein monofilaments and give sufficient details on the methods and results for analysis.[7,9–12] However, none of those studies have assessed changes in relative humidity, and only one of them included calibration of the filaments.[7]

A critical step in designing a study of sensory detection threshold, whatever the sensory modality, is the choice of a protocol. Sensory detection is not an "all-or-none" phenomenon. Therefore, the threshold is usually defined as the stimulus intensity at which the signal is detected 50% of the time.[13,14] A "gold standard" procedure for threshold assessment is the method of constant stimuli in which five or more different stimuli are presented at least 20 times each, in random order, to estimate the probability of detection of each stimulus (see Appendix A for a detailed description and discussion of the protocols). This method is time-consuming, as is the forced-choice task recommended for research studies by the Peripheral Neuropathy Association.[6] As psychophysical methods rely on a subject's cooperation and vigilance, the long duration of a test may be a source of inaccurate responses, especially when several sites are to be tested. Much faster procedures exist, the most popular being the method of limits. This method is subject to response bias, for example, anticipation bias. An alternative to these procedures is the family of adaptive procedures, in which stimulus intensity depends on the subject's response to the previous stimulus. Dyck et al.[15] described a "4, 2, 1" adaptive algorithm, in which the step between two consecutive stimuli is initially large and then progressively reduced, to allow the investigator to reach intensity values surrounding the sensory threshold rapidly. This procedure reduces the duration of the experiment and the subject's fatigue.

Abbreviations: ANOVA, analysis of variance; QST, quantitative sensory testing

Key words: adaptive staircase, quantitative sensory testing, Semmes–Weinstein monofilaments, tactile sensitivity, touch–pressure threshold

Correspondence to: A. D. Berquin; e-mail: anne.berquin@uclouvain.be

Published online 15 August 2010 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/mus.21689

This algorithm has been shown to give results comparable to those from a forced-choice algorithm, with substantial time-saving.[15] However, the alternation of ascending and descending series may still be perceived by the subject, thus possibly generating anticipation bias.

The aims of our study were to: (1) assess the reliability of Semmes–Weinstein monofilaments for detection of touch–pressure thresholds; (2) develop and validate a procedure suitable for rapid and accurate measurement of touch–pressure detection thresholds; and (3) establish normative data in healthy subjects, with special reference to age, gender, and body region.

METHODS

Materials and Calibration. The same set of 17 nylon monofilaments (Semmes–Weinstein Aesthesiometer; Senselab), whose forces, as indicated by the manufacturer, to range from 5.04 g/mm^2 (57 mg) for the smallest filament (number 3) to 178.25 g/mm^2 (140 g) for the largest filament (number 19), were used throughout the entire study. By manufacture, the bending force grows approximately exponentially with the increment in filament numbering. As shown in what follows, in normal subjects, thresholds were usually found between filament numbers 5 and 10. In some patients, however, the largest filament could hardly be detected.

Two calibration experiments were performed: (1) measure of the force applied on a balance by each filament, at a fixed humidity level; and (2) measure of the influence of the humidity level on the force exerted by a subset of four filaments (see Appendix B).

Standard Procedure for Touch–Pressure Threshold Measurements. The subjects were comfortably seated at a table (for measurements in the arms) or lying on a bed (for measurements in the legs), in a quiet room. It was not always possible to avoid the presence of other persons (e.g., relatives) in the room; this was accepted, as our procedure is designed for use at the bedside and not the laboratory. Any person present in the room during testing was asked to be as quiet as possible.

The subjects were informed that we wanted to measure their cutaneous sensitivity. The series of filaments was shown, and one large and one small filament were openly tested on their hand, so that they could see and feel a difference. Usually, they could not detect the smallest filament. We then explained that we would test different filaments in a sequence calculated by the computer, and we asked that they carefully concentrate on the tested site and determine, for each filament, if they felt the stimulus or not.

The computer screen and the box containing the monofilaments were placed out of sight of the subject. For measurements in the legs, the position of the subject was such that he could not see the tested site. For arm measurements, the subject was asked to close his eyes (no blinding device was used). If necessary, the cutaneous site to be tested was gently shaved to avoid stimulating skin hairs with the filaments. Room temperature and humidity were recorded before each threshold measurement. The site to be tested was indicated to the subject by a gentle touch of the investigators's hand.

The monofilaments were approached to approximately 3 cm of the tested site, oriented perpendicularly to the skin, and slowly brought into contact with the skin, taking care to avoid rapid impact or lateral movements. The filaments were then bent to about three-fourths of their extended length, for about 1 s. The investigator asked, "Do you feel this?" and then slowly removed the filament (again carefully avoiding lateral movements). The subject's answer was recorded on a computer (0 = not detected, 1 = detected). The computer calculated the intensity of the next stimulus, depending on the algorithm used (see later). Blank stimuli were applied by turning the handle of the filament by 90°, so that the filament was parallel to the skin but the subject perceived the same investigator movements as usual. At the end of the testing, the sequence of stimuli and subject's answers were displayed on a computer screen (as in Fig. 1b), to allow visual inspection of the results. All data were saved for subsequent analysis.

The following testing sites were used: for the upper extremities, we tested the first dorsal interosseous space of the hand (subsequently referred to as "C6") and the medial aspect of the forearm, about 5 cm distally to the elbow ("C8T1"); for the lower extremities, we tested the medial aspect of the lower leg, about 10 cm distal to the knee ("L4"), the first dorsal interosseous space of the foot ("L5"), and the external edge of the foot ("S1"). Subjects were tested either on the lower extremity or on the upper extremity. Sites were assessed in random order to average out the effects of fatigue.

All measurements were done in accordance with the Helsinki Declaration of 1975.

Stepping Algorithm. To avoid anticipation response bias that could occur with the original "4, 2, 1" staircase algorithm of Dyck et al.,[15] we interlaced two such staircases, beginning at a high and a low intensity level, respectively (filament

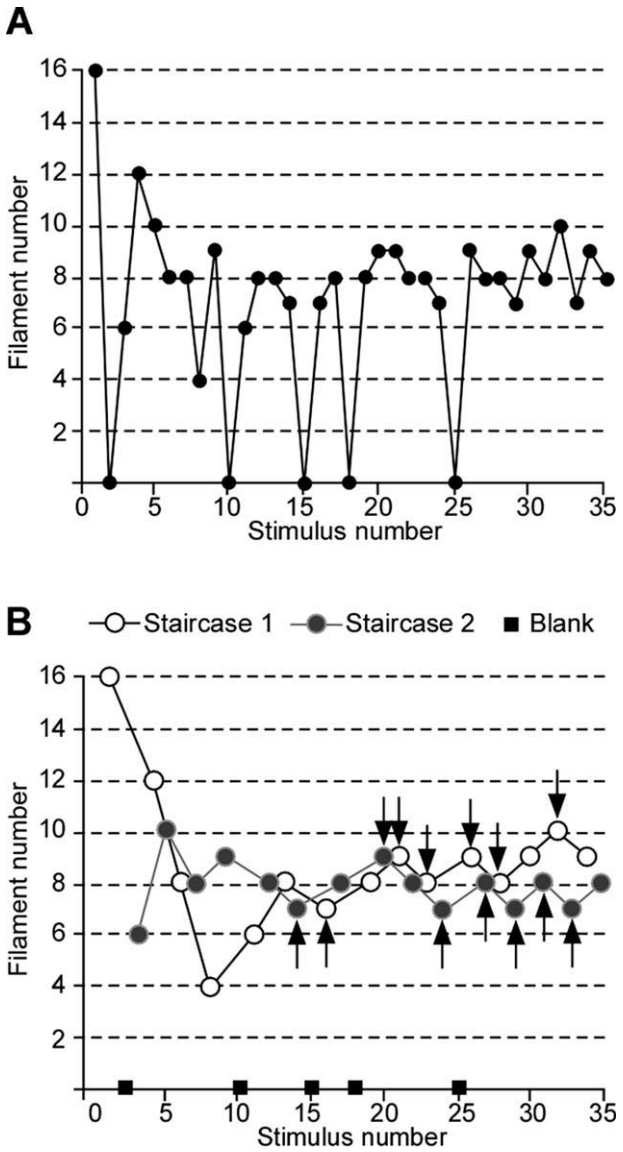

FIGURE 1. Results of an experimental session. **(A)** Series of stimuli as presented to the subject. **(B)** The same series of stimuli showing the interlacing of the two staircases. The arrows show the turnaround points averaged for threshold calculation.

numbers 16 and 6). Thus, odd-numbered stimuli belonged to the first staircase, whereas even-numbered stimuli belonged to the second staircase. A total of five blank stimuli were inserted at random (one of the blank stimuli was included in the first three stimuli, as done by Dyck et al.[15]). The total number of stimuli was thus 35. The value of "a" corresponded to one unit in filament numbering (as the force difference between two consecutive monofilaments is not linear but exponential, the interval a is not constant when expressed in milligrams). This procedure was computerized using an Excel spreadsheet.

The touch–pressure detection threshold was calculated by averaging all turnaround points obtained when $a = 1$ (thus excluding data obtained at the beginning of the experiment), after calculation of the force and pressure exerted by each filament, using filament numbering, calibration data, and the humidity measured at the beginning of the experiment. To correct for humidity changes, an equation was derived from the nomograph supplied by the manufacturer (see Appendix B). The percentage of false positives (i.e., detection of blank stimuli) was also calculated.

The validity of each measurement was assessed in two different ways. First, visual inspection of the experimental data ensured that the two staircases converged "reasonably well" to a plateau value. Second, data were rejected if the percentage of false positives was higher than 20% (i.e., if the subject "detected" more than one of five blank stimuli).

Methods of Limits and Constant Stimuli. The method of limits was used for gross evaluation of the threshold before testing for the method of constant stimuli. We used six series of stimuli, alternating ascending and descending series. The threshold was computed as the average of the pressure generated by the first filament detected in ascending series and the first filament not detected in descending series.

In the method of constant stimuli, five filaments surrounding the sensory threshold (previously estimated by the method of limits as just explained) were selected. These five filaments were presented 20 times each, in a random-order series (the total number of stimuli was thus 100). For each of these five filaments, a probability of detection was calculated and transformed to logit values. A linear regression was then used to calculate the sensory threshold, defined as the intensity of stimulation with a probability of detection of 0.5.

Validation of the Staircase Algorithm. The staircase procedure was validated against the method of constant stimuli, used as a "gold standard." In order to test the algorithm within a large range of threshold values, we selected 13 subjects suffering from known neurological diseases that reduce tactile sensitivity in either one or both lower limbs (sciatic lesions, multiple sclerosis, medullary pathology, diabetic polyneuropathy). One subject had to be excluded from the analysis because of intense paresthesia that hampered stimulus detection. Data obtained from the 12 remaining subjects (8 men, 4 women, mean age 52 ± 13 years) were used for the analysis. The body site to be tested (either L5 or S1) and the first side assessed (left or right) were randomly selected. To average out the effects of fatigue, measurements were performed in a mirror design; that is, if in the first site the staircase procedure was performed before the

constant stimuli method, then the constant stimuli algorithm was applied first at the contralateral site, and reciprocally.

Establishment of Normative Data. Healthy subjects were recruited among colleagues and relatives of two of our investigators (A.D.B. and V.L.) and partly via announcements posted in our hospitals. Exclusion criteria were all neurological diseases that could alter tactile sensitivity, such as sciatica (defined as low back pain irradiating below the knee), polyneuropathy, or diabetes. Subjects with a history of low back pain were accepted only if the pain never radiated below the knee. A clinical examination was performed, including muscle testing (biceps, triceps, and palmar interossei for upper limbs, and quadriceps, tibialis anterior, and triceps surae for lower limbs), clinical evaluation of tactile sensitivity (C6, C7, and C8–T1 dermatomes for upper limbs, and L4, L5, and S1 for lower limbs), and observation of osteotendinous reflexes. Subjects would have been excluded if any abnormality was observed (this did not occur). For each subject, date of birth, gender, height, weight, and self-reported laterality (right- or left-handed) were recorded. Measurements in the upper and lower limbs were performed by two different investigators (V.L. for upper limbs, A.D.B. for lower limbs); V.L. was trained by A.D.B. to ensure standardization of the procedure.

For measurements in the lower limbs, 4 male and 4 female subjects were recruited in each of the following age categories: 20–29 years; 30–39 years; 40–49 years; 50–59 years; and 60–69 years. The total number of control subjects was thus 40. As the results showed a slight but statistically non-significant trend for an increase in tactile sensitivity threshold with age, we decided to select control subjects for measurements in the upper limbs in only two age categories: 20–29 and 50–59 years (10 males and 10 females in each group).

Statistical Analysis. Throughout this study, data are presented as mean ± standard deviation. Homogeneity of demographic continuous variables between upper and lower limb groups was assessed by two-tailed unpaired *t*-tests.

Repeated measures analysis of variance (ANOVA) was performed on touch–pressure thresholds for the upper and lower limbs separately, with body site (L4, L5, and S1 or C6 and C8–D1) and laterality (right and left) as within-subject factors and age and gender as between-subject factors. When specific questions were raised about the effects of a subset of the levels of factors, a contrast analysis was performed. Two-tailed $P < 0.05$ was considered significant.

RESULTS

Reliability of Semmes–Weinstein Monofilaments for Detection of Touch–Pressure Thresholds. Detailed results of the calibration experiments are given in Appendix B. They show a poor relationship between manufacturer- and investigator-produced calibrations; however, the agreement between thresholds based on both calibrations was good, probably because of averaging effects. Changes in relative humidity, as observed in our experiments, result in large variations of the forces exerted by the filaments. For example, within the humidity range of 40–67% measured in our experiments, the pressure exerted by filament number 8 (close to the sensory threshold in most subjects), varied from 11.05 to 4.38 g/mm^2. The agreement between force and pressure values was acceptable in the physiological range, but it decreased with increasing threshold values. As a consequence, we decided to base further analysis on the pressure data obtained after correction for humidity and calibration results.

Validation of the Staircase Algorithm. Figure 1 shows an example of data obtained during one experimental session. It can be seen that the interlacing of the two staircases, as seen from the subject's standpoint (Fig. 1A), resulted in unpredictable changes in stimulus intensity. This impression was confirmed by some subjects who knew the algorithm but spontaneously remarked that, after a few stimuli, it was impossible for them to keep track of which was the next step in stimulus intensity. Another observation from Figure 1 is that, as expected, data converge very well to a plateau value corresponding to the sensory threshold.

When the results obtained with the staircase algorithm were compared with those obtained with the method of constant stimuli, relatively good agreement was obtained. The linear regression computed with these data indicated a slope of 1.045 and a negligible intercept of 0.46 ($r^2 = 0.90$). By contrast, the regression effective calculated when comparing the method of limits with the method of constant stimuli had a slope of 0.755 ($r^2 = 0.94$), indicating that the method of limits tended to underestimate high thresholds.

The average time needed to measure one threshold was 5 minutes for the adaptive procedure and the method of limits and 21 minutes for the method of constant stimuli, indicating a substantial time-saving for the adaptive procedure.

Establishment of Normative Data. Figure 2 shows a box-and-whisker plot for touch–pressure thresholds measured in lower and upper limbs. In the lower limbs, data are more scattered and thresholds seem to be higher in males than in females. In

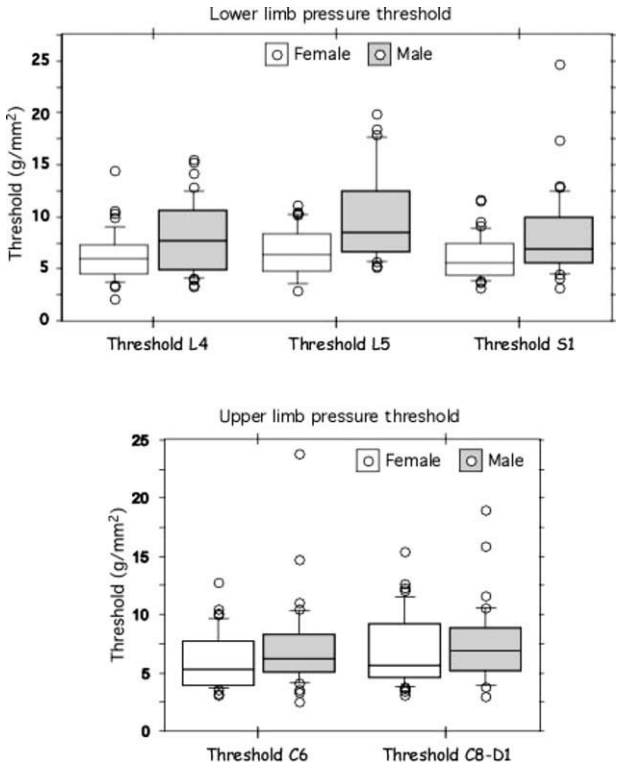

FIGURE 2. Box-and-whisker plots for touch–pressure thresholds.

upper limbs, gender differences seem to be less pronounced than in lower limbs. In both limbs, small differences between body sites could be observed.

In the lower limbs, repeated measures ANOVA for touch–pressure thresholds showed statistically significant differences according to gender ($P = 0.0021$) and site ($P = 0.012$); significant site × laterality ($P = 0.0025$) and site × laterality × gender ($P = 0.049$) interactions emerged. No significant influence of age was observed. Average thresholds ranged from 5.9 ± 1.9 g/mm^2 (L5 left, females) to 10.3 ± 4.5 g/mm^2 (L5 right, males).

In the upper limbs, repeated measures ANOVA showed that age was the only variable that influenced threshold measurements significantly ($P = 0.031$). The average threshold was 6.0 ± 2.9 g/mm^2 in young subjects and 7.7 ± 3.2 g/mm^2 in old subjects.

When the lower limbs were tested against upper limbs, no statistically significant difference was observed.

The average of absolute right/left differences ranged from 1.6 ± 2.2 to 2.7 ± 2.5 g/mm^2, depending on the body site tested.

Table 1 gives reference values for tactile sensitivity thresholds. As observed by Rolke et al.,[3] the frequency distribution of threshold values was skewed when expressed in the raw data and was closer to a normal distribution after logarithmic transformation (data not shown). Therefore, the 95% confidence interval was calculated from the logarithmic data and later back-transformed to non-logarithmic values. This explains why this interval is asymmetrical with respect to the mean.

Table 1. Mean, standard deviation, and 95% confidence intervals for tactile sensitivity touch–pressure thresholds.

			Pressure				Force	
Site			g/mm^{2*}	log g/mm^{2*}	g/mm^{2†}	g/mm^{2‡}	mg*	mg†
Upper limbs	Young		6.0 (2.9)	0.74 (0.18)	5.0–6.0	2.4–12.59	214 (221)	131–183
	Old		7.7 (3.2)	0.86 (0.16)	6.7–7.9	3.47–15.14	342 (309)	233–312
	Females	C6	6.0 (2.4)	0.75 (0.17)	5.0–6.3	2.57–12.3	207 (160)	128–199
		C8T1	6.8 (3.1)	0.79 (0.18)	5.4–7.0	2.69–14.13	267 (225)	151–249
		All	6.4 (2.8)	0.77 (0.17)	5.4–6.4	2.69–12.88	237 (196)	149–208
	Males	C6	7.1 (3.7)	0.81 (0.18)	5.7–7.3	2.82–14.79	314 (386)	173–282
		C8T1	7.4 (3.2)	0.84 (0.17)	6.1–7.8	3.16–15.14	325 (275)	203–315
		All	7.3 (3.4)	0.83 (0.18)	6.2–7.4	2.95–15.49	319 (333)	201–279
Lower limbs	Females	L4	6.2 (2.4)	0.76 (0.16)	5.1–6.5	2.75–12.02	230 (161)	152–230
		L5	6.5 (2.4)	0.78 (0.17)	5.3–6.8	2.75–13.18	261 (167)	165–260
		S1	6.0 (2.1)	0.76 (0.14)	5.2–6.4	3.02–10.96	216 (137)	145–216
		All	6.2 (2.3)	0.77 (0.16)	5.5–6.3	2.82–12.3	236 (155)	168–215
	Males	L4	8.1 (3.4)	0.87 (0.19)	6.5–8.5	3.09–17.78	376 (283)	219–362
		L5	10.1 (4.5)	0.97 (0.18)	8.2–10.6	4.07–21.38	598 (493)	349–562
		S1	8.1 (4.1)	0.87 (0.19)	6.5–8.5	3.09–17.78	388 (377)	222–361
		All	8.8 (4.1)	0.90 (0.19)	7.3–8.6	3.31–19.05	454 (403)	284–379
All data	Females		**6.3 (2.5)**	**0.77 (0.16)**	**5.6–6.2**	**2.82–12.3**	**236 (172)**	**167–203**
	Males		**8.2 (3.9)**	**0.87 (0.19)**	**7.0–7.9**	**3.09–17.78**	**400 (382)**	**258–322**

Pressure values obtained from investigator-produced calibrations and taking hygrometry into account; force values obtained from investigator-produced calibrations and taking hygrometry into account. Bold values indicate global normative data. SD, standard deviation; CI, confidence interval.
**Data expressed as mean (SD).*
†Data expressed as 95% CI.
‡Data expressed as mean ± 2 SD.

Force values are given for a better comparison with other published studies.

DISCUSSION

Quantitative sensory testing is useful for assessing alterations in sensory modalities resulting from physiological and pathological conditions. However, its clinical and research applications are hampered by the lack of standardization and normative data. Moreover, one of the most widely used devices for mechanical sensitivity assessment (Semmes–Weinstein monofilaments) is subject to criticism. This study was designed to: (1) assess the reliability of Semmes–Weinstein monofilaments for detection of touch–pressure threshold; (2) develop and validate a procedure suitable for rapid and accurate measurement of touch–pressure detection thresholds; and (3) establish normative data in healthy subjects, with special reference to age, gender, and body region.

Reliability of Semmes–Weinstein Monofilaments for Measure of Touch–Pressure Thresholds. As mentioned earlier, several caveats undermine the use of Semmes–Weinstein monofilaments for tactile sensitivity threshold measurements.[6–8,16] However, these filaments are still widely in use, probably because of their simplicity and relatively low cost when compared with more sophisticated systems. An alternative to the classical Semmes–Weinstein filaments is optical glass fibers, which are not sensitive to changes in relative humidity and can be equipped with a rounded tip of standardized size.[17] However, these fibers are not yet widely available.

This study confirms that the use of Semmes–Weinstein filaments is not as straightforward as is usually assumed by investigators. Calibration experiments must be performed at least once with each set of filaments. Humidity changes must be taken into account, and the procedure of filament application must be standardized as much as possible. This observation is at odds with the lack of calibration and humidity measurements widely observed in the literature.

Validation of the Staircase Algorithm. Psychophysical tests are subjective in nature, as they rely on the subject's responses, which may be influenced by both conscious and unconscious factors. Therefore, their validity is subject to a number of potential pitfalls.[1] First, they are critically influenced by subject cooperation and may therefore be influenced by boredom, mental fatigue, or distraction. Thus, algorithms of short duration, such as our adaptive algorithm, are preferable to longer lasting procedures. Second, unconscious anticipation bias can occur when the sequence of stimuli is such that the subject, after a few stimuli, becomes able to guess the intensity of the next stimulus. This is of particular concern with the method of limits. In the algorithm presented here, the interlacing of two staircases is such that the subject is unable to recognize the rules determining the sequence of stimuli. Third, other unconscious factors such as anxiety, as well as sensory factors such as paresthesia, may hamper the discrimination of stimuli against background sensory noise. The use of blank stimuli allows us to test for this possibility and reject data if more than one of five blank stimuli can be described as positive. Finally, conscious bias such as malingering can influence the results. Although there is no definitive method to overcome this problem, two aspects of our algorithm can help detect unreliable responses: (1) the interlacing of two staircases, as discussed earlier; and (2) the assessment of data convergence to a plateau, which is a condition for considering our adaptive algorithm valid. The assumption that the stochastic process of detection is stationary is indeed critical for the procedure to be applicable.[18]

The stepping algorithm described in this study, when compared with the method of limits and the method of constant stimuli (considered as a reference), showed better accuracy than the method of limits. There was substantial time-saving compared with the method of constant stimuli. The reproducibility of our algorithm has not been tested, but filament reproductibility has been considered acceptable by other investigators.[19] Moreover, as suggested by Rolke et al.,[3] the good correlation of data obtained on the right and left sides of the body shows good reproducibility in the short term.

Establishment of Normative Data. The literature reporting tactile sensitivity thresholds in physiological and pathological conditions is abundant. However, we found only a few studies that gave sufficient details on their methods and results to permit effective comparison with our data.[3,7,9–12] These studies are summarized in Appendix C (Table AC1). The experimental methods as well as the tested sites vary widely across studies. Only one study[7] relied on filament calibration, and no study assessed for humidity. The number of subjects was small in most studies, except for that of Rolke et al.[12] The age range was usually wide. As the units of threshold measurements were very different, comparisons of threshold values are difficult. We have tried to convert all data to force values expressed in milligrams (as filament cross-sectional areas were not available, the calculation of pressure values from data in the literature was impossible). However, several studies did not give enough information to ensure that the conversion was

valid. Threshold values usually lie between 10 and 700 mg. Differences between the results could be due to numerous factors, such as algorithms, materials, lack of calibration and measure of humidity, body sites, and subjects. This once again stresses the importance of establishing reference values in each laboratory.

In good accordance with our data, gender differences showed a trend for lower threshold values in females, but differences were small and not always statistically significant. However, our data are more scattered in males than in females for the lower limb, and therefore the range of normal values is wider in males than in females.

Most studies have shown a threshold increase with age, and the magnitude of this effect was very different depending on the study. However, the age-dependent reduction in tactile sensitivity was most prominent in older subjects (>68 years of age).[9,11] This could explain why the age effect was small in our study (our oldest subject was 69 years old). The decrease in tactile sensitivity with age may be correlated with a decrease in the density of Meissner corpuscles.[9]

Results concerning site differences have been mixed. In three studies,[3,11,12] a craniocaudal decrease in tactile sensitivity was observed; Halar et al.[10] observed lower thresholds in proximal than in distal areas in lower limbs. Our study and that of Voerman et al.[7] did not find significant differences in the upper limbs, whereas we found statistically significant differences (however small in absolute value) between L4, L5, and S1 sites in the lower limbs, and no differences in average values for upper vs. lower limbs. It must be stressed that all studies, except that by Rolke et al.,[12] used small numbers of subjects, which may have hampered the detection of small differences. One other factor that may contribute to the lack of differences between upper and lower limbs in our study is that measurements were performed by different examiners. Although V.D.L. was trained by A.D.B. to ensure standardization of the method, small differences in the protocols (e.g., filament handling) could have occurred and may have undermined the validity of any upper vs. lower limb comparison.

Finally, right/left differences were absent in all studies where this parameter was assessed. This suggests that, in areas for which no normative data are available, the contralateral site might be used as a control. Moreover, given that the interindividual differences are higher than the right–left differences, right–left comparisons might be more sensitive to small sensory alterations than comparison with absolute reference data.[12] However, this is not valid in symmetric pathologies such as polyneuropathies.

To summarize, our data and several other studies suggest that body site (below the elbow for the upper limb and below the knee for the lower limb), age (<70 years), or laterality have only a marginal effect on threshold measurements. Therefore, we suggest using the global normative data indicated in bold in Table 1 for routine clinical assessment. For research purposes, however, small but statistically significant differences may be of importance, and more detailed normative data, taking age, gender, and body-site differences into account, would be necessary. Given the fact that the results are highly influenced by the experimental setting (standardization of instructions to subjects, training of technicians, machine calibration, stimulus characteristics, and testing algorithms), each research laboratory must determine its own set of normal values.[1]

Conclusion. As a first step before gathering data from patients, we have come to three conclusions in this study. First, an analysis of the reliability of Semmes–Weinstein monofilaments showed that each set of monofilaments should be carefully calibrated at least once, and that humidity should be routinely assessed at the beginning of each experiment, as it can cause a more than twofold change in pressure values. Taking humidity changes into account can be easily implemented in any spreadsheet program. Second, validation of the adaptive algorithm shows good accuracy and substantial time-saving when compared with the method of constant stimuli. Finally, measures in normal subjects suggest that, within the limits of our experimental data (distal parts of the limbs and age <70 years), body site, age, and laterality have only a small influence on mechanical sensitivity thresholds, whereas gender differences are observed in the lower limbs. Therefore, for most clinical purposes, general reference values may be used. In contrast, any research work must rely on normative data carefully established in each research setting.

APPENDIX A: ALGORITHMS FOR THRESHOLD CALCULATION

In perceptual decision processes, there is no sharp transition from a non-detect to a detect state. That transition is best described by the probability of detection as a function of stimulus intensity and depicted as an S-shaped curve. Therefore, the threshold is usually defined as the stimulus intensity at which the stimulus is detected 50% of the time. Several methods have been described to measure sensitivity thresholds.

The "gold standard" is the method of constant stimuli, in which five or more intensity levels are presented at least 20 times each, in a random-

order series. The probability of detection of each stimulus intensity can be calculated, allowing the investigator to determine the psychophysical function (probability of detection vs. stimulus intensity) and to calculate the stimulus intensity corresponding to a probability of detection of 0.5.[13,14] This method is time-consuming and therefore the subject's cooperation may decrease with time, thus affecting the results. Moreover, it critically relies on the necessity to choose a priori a range of stimuli centered on the presumed threshold, which can be problematic in pathological conditions.

A much faster and very popular procedure is the method of limits. Stimulus intensity is presented in ascending or descending series, and the threshold is calculated as the geometric mean of several series. However, this method is highly subject to response bias, which may be due to unconscious cognitive factors, such as anticipation bias,[13] more often than malingering.

In the family of adaptive procedures, stimulus intensity is basically a function of the subject's response to the preceding stimulus: if a stimulus S_n is perceived, then the intensity of the next stimulus is decreased by a step "a" ($S_{n+1} = S_n - a$); if S_n is not perceived, the intensity of the next stimulus is increased by the same step ($S_{n+1} = S_n + a$). "Turnaround points" designate the points where the direction of intensity changes is reversed (when a descending series is followed by an ascending series, or vice-versa). The step "a" between two consecutive stimuli may be of a fixed value or may be progressively decreased. In the "4, 2, 1" adaptive algorithm of Dyck et al.,[15] the step value is initially set to $4a$ and is reduced to $2a$ and, finally, to a after the first and second turnaround points, respectively. This speeds up the initial part of the procedure, thus reducing subject fatigue.

APPENDIX B: CALIBRATION METHODS AND RESULTS

Methods. ***Calibration 1 Experiments.*** The filaments were calibrated by measuring the force applied at the plateau of a balance (Mettler Toledo College B154 for filaments 3–12, Mettler K7 for filaments 13–19). Each filament was slowly applied perpendicularly to the plateau and bent to approximately three-fourths of its extended length; the applied force was recorded in milligrams or grams. Six consecutive measurements were made for each filament, and the average of measurements 2–6 was computed.

Calibration 2 Experiments. Information supplied with the filament set stated that the force required to buckle the nylon monofilament varies according to humidity. A nomograph was given, showing the relationship between log force and humidity, for each filament. We used this nomograph to establish, for each filament, an equation computing the force applied at a relative humidity of x%: $F_x = 10\mathrm{E}[\log(F_{25}) + (A_n{*}(x - 25)]$, where F_x is the force at relative humidity x, F_{25} is the force at 25% humidity (as measured from the nomograph), and A_n is the regression coefficient for the filament n (estimated from the nomograph). This equation was experimentally tested with four filaments by measuring the force applied by these filaments on a balance, as explained previously, at two different humidity levels (19% and 28%; humidity was increased in the laboratory by boiling water in an open pan). The results obtained in the Calibration 1 experiment, on a different day, at a humidity level of 34%, were added to the data.

Force Calibration Results. When the forces measured on a balance (at a humidity level of 34%) were compared to manufacturer-produced calibrations (relative humidity not specified), the differences between measured and theoretical data ranged from −19.75% to 17.61%, indicating wide variations. The disagreement between measured and theoretical values could be either positive or negative, with no systematic pattern of variation. This suggests that the disagreement cannot be accounted for by differences in relative humidity.

In contrast, the agreement between threshold values calculated with manufacturer-produced calibrations, and those calculated with investigator-produced calibrations was, surprisingly, very good (regression coefficient = 1.02, intercept = 0.04, r^2 = 0.98). One explanation for this may be that the discrepancy between manufacturer- and investigator-produced calibrations was cancelled out by the averaging process: threshold was calculated by averaging the force (or pressure) for all turnaround points. In a typical experiment, the 10–18 turnaround points distributed themselves to three or four adjacent filaments, mainly among filaments numbered 5–10, in which the difference between manufacturer- and investigator-produced calibrations was alternatively negative and positive.

Relative Humidity Calibration Results. When measurements taken at different humidity levels were compared with manufacturer-produced calibrations (data not shown), the agreement was found to be relatively good, except for one data point, which was excluded from further analysis. However, the relative difference between measured and calculated data could still be as high as 35%.

Because the humidity measured during the experimental sessions ranged from 40% to 67%, one can expect that it would greatly influence our results. For instance, at these two relative humidity values, the pressure exerted by filament number 8 (closest to the sensory threshold in most subjects),

varied from 13.1 to 5.19 g/mm^2, according to manufacturer-produced calibrations, or from 11.05 to 4.38 g/mm^2, according to investigator-produced calibrations. As a consequence, neglecting the humidity effect systematically alters threshold evaluation (the magnitude of this overestimation being higher at high threshold values). Therefore, data analysis was performed taking relative humidity into account.

Units of Measurement. The relationship between threshold values expressed in force units (mg) versus the values expressed in pressure units (g/mm^2) was approximately linear for threshold values below 12 g/mm^2; above this value, the slope of the relationship increased (i.e., force data increased faster than pressure data; data not shown). As the average normal threshold is close to 7 g/mm^2, this discrepancy will be more marked in patients with hypoesthesia than in normal subjects. Moreover, data dispersion was higher with force results than with pressure results. For these reasons, as well as physiological reasons, we decided to use pressure values for further calculations, but kept force values for comparisons with the results of other investigators.

APPENDIX C

Table AC1. Published data concerning tactile sensory threshold with respect to gender, age, site, and laterality.

	Thornbury and Mistretta[9]	Halar et al.[10]	Investigators Voerman et al.[7]	Haanpaa et al.[11]	Rolke et al.[3,12]	This study
Method	Stepping adaptive algorithm (step = 1)	Forced choice	3 applications of each filament (order not specified), threshold = force of smallest filament detected at least 2 times	Limits (number of series not specified)	Limits (10 series, alternating ascending and descending, geometric mean)	Stepping adaptive algorithm (4, 2, 1 steps)
Correction for hygrometry	no	no	no	no	no	yes
Filament calibration	no	no	yes	no	no	yes
Number of male subjects	24	23	10	30	9 (a), 70 (b)	20 upper limb, 20 lower limb
Number of female subjects	31	13	10	73	9 (a), 110 (b)	20 upper limb, 20 lower limb
Age range	19–88 years, 4 groups (means 31, 47, 66, and 79 years)	21–77 years, 6 groups (<29, 30–39, 40–49, 50–59, 60–69, 70–80)	Mean 34.7 years (SD 8.8)	19–90 years (mean 65) 4 groups (<60, 60–69, 70–79, >80)	17–75 years, mean 38.4 (SD 12.9) (b)	20–70 years 5 groups (20–29, 30–39, 40–49, 50-59, 60–69) for lower limb, 2 groups (20–30 and 50–60 for upper limb)
Site(s)	Volar side distal phalanx dominant index finger	12 sites upper extremity, 10 sites lower extremity, non-dominant side	C3 to C8	V1, C3, T3, T10, S1	Cheek, hand dorsum, foot dorsum	C6, C8T1, L4, L5, S1
Unit of threshold measure	Log10 0.1 of mg	Log10 of 0.1 mg of force	Log10 of 0.1 mg of force	Authors state Log10 of mg, but data suggest Log10 of 0.1 mg	mN	g/mm^2
Published threshold range (min–max values)	2.18–2.83	1.97–3.27	2.34–3.86	3.22–5.88	0.21–3.52 (a)	5.5–10

Table AC1. (Continued).

	Thornbury and Mistretta[9]	Halar et al.[10]	Investigators Voerman et al.[7]	Haanpaa et al.[11]	Rolke et al.[3,12]	This study
Threshold range (mg)	15.14–67.61	9.33–186.21	21.88–724.44	165.96–8511.38? (unclear units of measure)	21.4–358.8 Face 21.4, SD 5.1; hand 196.7, SD 212; foot 358.8, SD 352 (a)	176–579
Gender differences	Lower threshold in females but not statistically significant	None	None	Lower threshold in females only in thoracic dermatomes	None	Lower threshold in females, small differences, statistically significant in lower limbs
Age differences	Threshold increases with age: 15 mg in youngest, 68 mg in oldest (most prominent above 66 years)	Gradual and progressive increase with age at 5 of 12 sites in upper extremity and 3 of 9 in lower extremity (size effect not available)	Not assessed	Threshold increases with age, mainly >70 years; >80 years males, mean threshold 75.8 g	Threshold increases with age (detailed data not available) (b)	Old > young, statistically significant in upper limbs only (young 6 g/mm^2, old 7.7 g/mm^2)
Site differences in measured thresholds	Not assessed	Upper limb (mean 17 mg) < lower limb (mean 110 mg), $P <$ 0.005. In upper limb proximal = distal, in lower limb proximal < distal	None	C3 to S1: Cranial < caudal (statistics not available)	Face (21.4 mg) < hand (196.7 mg) < foot (358.8 mg) (a)	None or small
Right–left differences	Not assessed	Not assessed	None	None	None	None
Remarks	33 subjects (mainly younger) can detect the smallest filament: threshold is overestimated in these subjects		Calibration showed large differences between expected ant measured forces; authors used calibrated data	Data conversion to mg of force is hazardous due to lack of details concerning the units used in the study	Study a[3] preliminary, 18 subjects; study b[12] multicenter study with standardized protocol	

SD, standard deviation.

A.B. was supported in part by a Mathilde Horlait-Dapsens grant and by a grant from Medtronic, Inc.

REFERENCES

1. Chong PS, Cros DP. Technology literature review: quantitative sensory testing. Muscle Nerve 2004;29:734–747.
2. Rowbotham MC, Lottrup Petersen K, Fields HL. Is postherpetic neuralgia more than one disorder? Pain Forum 1998;7:231–237.
3. Rolke R, Magerl W, Andrews Campbell K, Schalber C, Caspari S, Birklein F, et al. Quantitative sensory testing: a comprehensive protocol for clinical trials. Eur J Pain 2006;10:77–88.
4. Yarnitsky D. Quantitative sensory testing. Muscle Nerve 1997;20: 198–204.
5. Shy ME, Frohman EM, So YT, Arezzo JC, Cornblath DR, Giuliani MJ, et al. Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Quantitative sensory testing: report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology 2003;60: 898–904.
6. Peripheral Neuropathy Association. Quantitative sensory testing: a consensus report from the Peripheral Neuropathy Association. Neurology 1993;43:1050–1052.
7. Voerman VF, van Egmond J, Crul BJ. Normal values for sensory thresholds in the cervical dermatomes. A critical note on the use of Semmes–Weinstein monofilaments. Am J Phys Med Rehabil 1999;78: 24–29.
8. Andrews K. The effect of changes in temperature humidity on the accuracy of von Frey hairs. J Neurosci Methods 1993;50:91–93.
9. Thornbury JM, Mistretta CM. Tactile sensitivity as a function of age. J Gerontol 1081;36:34–39.
10. Halar EM, Hammond MC, LaCava EC, Camann C, Ward J. Sensory perception threshold measurement: an evaluation of semiobjective testing devices. Arch Phys Med Rehabil 1987;68:499–507.
11. Haanpää M, Laippala PA, Nurmikki TJ. Thermal and tactile perception thresholds in acute herpes zoster. Eur J Pain 1999;3:375–386.

12. Rolke R, Baron R, Maier C, Tölle TR, Treede RD, Beyer A, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. Pain 2006;123:231–243.
13. Gescheider GA. Psychophysics methods, theory and application, 2. Hillsdale, NJ: Erlbaum; 1985.
14. Plaghki L, Mouraux A. Fonctions psychométriques et méthodes psychophysiques adaptatives pour l'étude de la douleur. Doul et Anal 2001;1:73–77.
15. Dyck PJ, O'Brien PC, Kosanke JL, Gillen DA, Karnes JL. A 4, 2 and 1 stepping algorithm for quick and accurate estimation of cutaneous sensation threshold. Neurology 1993;43:1508–1512.
16. van Vliet D, Novak CB, Mackinnon SE. Duration of contact time alters cutaneous pressure threshold measurements. Ann Plast Surg 1993;31:335–339.
17. Fruhstorfer H, Gross W, Selbmann O. von Frey hairs: new materials for a new design. Eur J Pain 2001;5:341–342.
18. Falmagne JC. Elements of psychophysical theory. New York: Oxford University Press; 1985.
19. Valk GD, de Sonnaville JJ, van Houtum WH, Heine RJ, van Eijk JT, 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;20:116–118.