AAEM PRACTICE TOPIC IN ELECTRODIAGNOSTIC MEDICINE

American Association of Electrodiagnostic Medicine
421 First Avenue S.W., Suite 300 East, Rochester, MN 55902 (507/288-0100)

ABSTRACT: The development of the personal computer has simplified the process of quantitating sensory thresholds using various testing algorithms. We reviewed the technical aspects and reproducibility of different methods to determine threshold for light touch-pressure, vibration, thermal, and pain stimuli. Clinical uses and limitations of quantitative sensory testing (QST) were also reviewed. QST is a reliable psychophysical test of large- and small-fiber sensory modalities. The results of QST are highly dependent on methodology and the full cooperation of the subject. QST has been shown to be reasonably reproducible over a period of days or weeks in normal subjects. The use of QST in research and patient care should be limited to instruments and their corresponding methodologies that have been shown to be reproducible. Literature data do not allow conclusions regarding the relative merits of individual QST instruments.

Muscle Nerve **29**: 734–747, 2004

TECHNOLOGY LITERATURE REVIEW: QUANTITATIVE SENSORY TESTING

PETER SIAO TICK CHONG, MD, and DIDIER P. CROS, MD

American Association of Electrodiagnostic Medicine, 421 First Avenue SW, Suite 300 East, Rochester, Minnesota 55902, USA

INTRODUCTION

The purpose of this technology literature review is to assess the methodology, reliability, reproducibility, limitations, and potential clinical applications of quantitative sensory testing (QST). This review discusses the use of QST in the perception of the following sensory modalities: (1) light touch; (2) vibration; (3) thermal; and (4) pain. Other sensory testing devices such as the current perception test, tactile circumferential discriminator, and two-point esthesiometer are not included in this review. No

clinical tests or trials were performed by the American Association of Electrodiagnostic Medicine (AAEM); the AAEM's opinion is based solely on a review of the literature.

OVERVIEW

In the past, clinicians and researchers have relied on clinical sensory examination and conventional sensory nerve conduction studies (NCSs) in the evaluation, management, and follow-up of patients with peripheral neuropathy (PN). Sensory NCSs are reliable and reproducible, especially when performed by a single examiner.[15,16,61] However, sensory NCSs can only evaluate large myelinated nerve fibers. For this reason, different tests are needed to evaluate small and unmyelinated nerve fibers.

QST can assess and quantify sensory nerve function noninvasively. The light-touch and vibration testing modalities evaluate the large myelinated A-alpha and A-beta sensory fibers, whereas the thermal testing modality assesses small myelinated and unmyelinated sensory nerve function.[76] In addition, confirmatory studies in primates and humans have shown that non-nociceptive cool stimuli are mediated by A-delta small myelinated fibers, and warm stimuli and nociceptive stimuli are mediated by C

5fibers.[2,19,20,22,35,40,46,63] The pain modality, heat-pain, and cold-pain threshold tests have been found to document thermal hyperalgesia and hypoaesthesia.[73] Therefore, QST can study the large myelinated, small myelinated, and unmyelinated fibers in addition to documenting hyperalgesia and hypoesthesia.

QST is now becoming more widely available because of increasing interest and technological improvements. QST has potential as a neurophysiologic tool. Using QST, sensory deficits may be quantified and the data can be used in parametric statistical analysis. The increasing acknowledgment of QST as a diagnostic tool is evidenced by the American Diabetes Association endorsement of QST in 1992 as a valid test in epidemiologic studies and drug trials of diabetic neuropathy.[5,6] In addition, QST is also being used in research studies and drug trials of other types of neuropathy to monitor sensory nerve function. Normal values have been established at a number of institutions. When establishing the normal values, researchers have found that age, sex, and the site of the stimulation can affect sensory thresholds.[28]

LITERATURE SEARCH

This technology review was based on searches of *Medline*, and references from relevant articles published between 1966 and 2001. The following search terms were used: "quantitative sensory testing"; "QST"; "sensory threshold"; "thermal threshold"; "temperature sense"; "vibration"; and "vibration threshold." The search for literature included only articles written in English.

TECHNICAL ASPECTS

The availability of automated systems has allowed the use of more detailed testing algorithms and has lessened some sources of error that were problematic with early forms of QST. Studies have found that standardization of instructions to subjects, training of technicians, machine calibration, stimulus characteristics, and testing algorithms are all essential for accurate and reproducible QST. To produce reliable results, the AAEM recommends that neither the technologist nor the patient should be aware of the previous results during follow-up studies. In addition, bilateral evaluations of several sensory modalities should be conducted to ensure a comprehensive evaluation of an individual patient because different sensory fibers may be preferentially affected and significant side-to-side asymmetry may be present.[78]

In contrast, unilateral testing may be adequate in longitudinal studies or population screening of a disorder such as diabetic polyneuropathy because it is usually symmetric.[5,6]

A number of different instruments and protocols have been described in the literature. Currently, there is a need to standardize testing procedures and data reporting to facilitate comparison of QST reports from different centers. The results of an individual patient can be compared with normative data (matched for age, sex, and site of stimulation) and expressed in a percentile.[62] The Peripheral Neuropathy Association recommends that only validated forced-choice testing algorithms should be used for research purposes.[62]

This article provides a discussion of the technical aspects of QST devices and methods of sensory threshold determination based on a review of the literature. It is the intent of this review to highlight technical differences. The review is not an endorsement by the AAEM of any product. It is an analysis of the available literature and is the opinion of the AAEM based on the literature.

METHODS OF THRESHOLD DETERMINATION

Sensory threshold testing is a psychophysical test. It requires that the subject is alert, cooperative, and able to follow instructions. The variance of sensory threshold measurements in normal subjects pretending to be abnormal is larger than that of trustworthy normal subjects and neuropathic subjects.[81] However, there is currently no reliable way to differentiate the results of subjects who are biased (whether consciously or subconsciously) toward an abnormal study from those with organic disorders.[25] As with most neurophysiologic methods, one cannot overemphasize the importance of deriving one's own set of normal values. Published normal values may be used if the equipment, methodology, and algorithm are identical to those used in establishing the normal values.

This review focuses on the major methods of threshold determination. The methods for light-touch, thermal, and vibration sensations are discussed together, and the method of heat-pain response is considered separately.

LIGHT-TOUCH, VIBRATION, AND THERMAL SENSATIONS

Sensory threshold determination using the light-touch, vibration, and thermal methods in QST involves the following three basic considerations that are interrelated and not mutually exclusive: (1) how

the stimulus is presented to the subject; (2) how the subject's response to the stimulus is obtained; and (3) how the subject's responses are utilized to determine sensory threshold. Each of these processes is discussed. It is important to note that different investigators use different combinations of these basic steps to determine the sensory threshold.
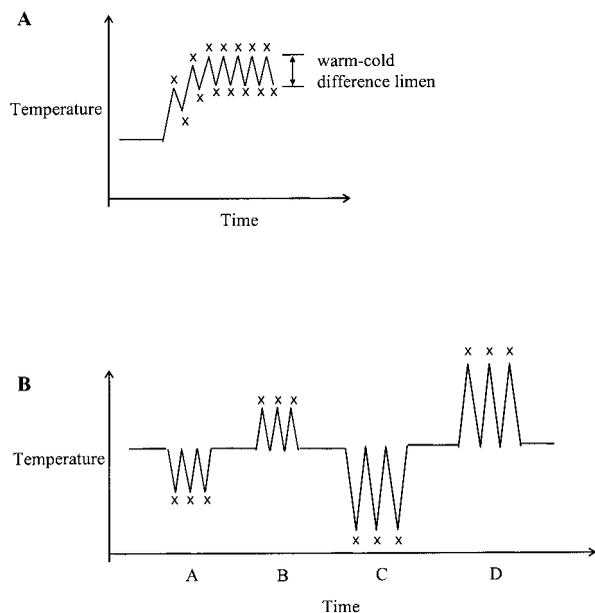
How the Stimulus Is Presented to the Subject. In any testing protocol, the characteristics of the stimulus should be specified and followed rigorously because any deviation may affect the result (i.e., sensory threshold determination). The model or make of the instrument, site of stimulus, size of stimulated area, ramp rate, and stimulus duration are all important variables to consider. The stimulus may be presented to the subject using two methods: (1) the method of limits; or (2) the method of levels.

Method of Limits. An example of the method of limits is the original Marstock method of determining thermal thresholds (Fig. 1A). This stimulation technique is similar to that in Bekesy audiometry.[36] In Bekesy audiometry, the tone increases in intensity

FIGURE 1. Method of limits (schematic diagram). **(A)** Marstock method: warm–cold difference limen is the term used by Fruhstorfer et al. to represent the indifferent temperature range in which no thermal sensations occur. The warm and cool thresholds are both reflected in the warm–cold difference limen. Abnormalities of either the warm or the cool threshold, or both, would result in a high warm–cold difference limen. **(B)** Modified Marstock method: cool threshold (A) and warm threshold (B) are measured separately. As with the original Marstock method, cold-pain threshold (C) and heat-pain threshold (D) are also measured.

until the subject perceives the sound and presses a button. The audiometer will then decrease the intensity of the tone until the subject can no longer hear the sound and presses the button again. To determine the thermal threshold, the patient also plays an active role. The thermal stimulator (thermode) is applied to the skin and the temperature of the thermode increases until the subject feels a warm sensation and gives a signal by pressing a button (shown in Fig. 1 as an X). This is the appearance threshold. This signal results in a reversal of the temperature change of the thermode (from warm to cool). The subject then gives another signal as soon as a cool sensation is perceived, termed the disappearance threshold. The second signal reverses the direction of temperature change again, from cool to warm, leading to another cycle. After several cycles (usually 1–2 minutes), the initial rises in thresholds secondary to adaptation stabilize, allowing the warm and cool thresholds to be measured together, as reflected in the warm–cold difference limen of the original Marstock method (Fig. 1A). In contrast, the modified Marstock method (Fig. 1B) measures the thresholds of the different sensory modalities separately (cool, warm, cold pain, and heat pain).[73] As soon as the subject perceives a warm or cool sensation, the button is pressed and the temperature of the thermode returns to baseline.

Inherent in the method of limits is the inclusion of reaction time, which results in overestimation of sensory threshold. The results tend to be variable because they depend on the subject's full cooperation and constant vigilance. Reaction time is dependent on factors such as concentration, drowsiness, or boredom, all of which are difficult to control. A "learning effect" with repeated testing is also possible.[80]

Other algorithms utilizing the method of limits include the linear ramp algorithm with null stimuli and the Bekesy algorithm with null stimuli.[27] In the linear ramp algorithm with null stimuli, the stimulus intensity increases linearly until the subject presses the response key (appearance threshold). Randomly occurring null stimuli are included with the true stimuli. The threshold is taken from the mean of appearance thresholds. The Bekesy algorithm with null stimuli is a modification of the Bekesy method for determining auditory thresholds. The subject presses the response key as soon as the stimulus is felt (appearance threshold) and releases the response key when the stimulus disappears (disappearance threshold). The rate of increase or decrease of the stimulus intensity follows an exponential function and not the $\log_{10}$ function of the Bekesy

method. The mean of the appearance and disappearance thresholds is taken as the threshold. Vibratory thresholds obtained from the Bekesy algorithm with null stimuli have been shown to be similar to those obtained with the forced-choice algorithm (discussed later). In contrast, the linear ramp algorithm overestimates the threshold.[27]

Method of Levels. The method of levels overcomes the disadvantage of the method of limits (inclusion of reaction time) by utilizing stimuli of predetermined levels of stimulus intensity and duration. When using this method, at the end of each stimulation the subject is asked whether the stimulus was perceived (Figs. 2 and 3). The success or failure of the subject's response then determines the next level of stimulus intensity (higher or lower). The reaction time is not included in the method of levels and overestimation or increased variability of sensory thresholds is minimized.

How to Obtain the Subject's Response to the Stimulus. A subject's sensory threshold is classically defined as the level of stimulus intensity necessary for a sensation to be just detectable. According to the signal detection theory, an external stimulus is detected if the subject decides that the level of neural activity is more than the level of "internal noise" or random activity. To determine the sensory threshold, the subject may be presented with a series of stimuli of different magnitudes and asked whenever a stimulus is detected. A stimulus is detected when the subject perceives an alteration of the baseline "internal noise." There are two methods of obtaining the response to the stimuli: the yes–no method and the forced-choice method.

Yes–No Method. In the yes–no method, the subject is presented with different stimuli of predetermined levels (method of levels) and is asked to respond to each with a "yes" or a "no"[80] (Fig. 2).
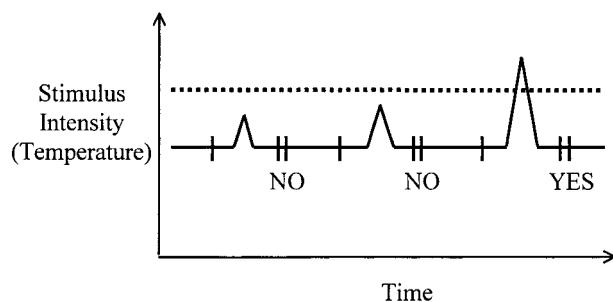
FIGURE 2. Method of levels with yes–no response. (|) Warning prior to stimulus; (⋀) stimulus of predetermined intensity; (‖) subject is asked to respond to whether a stimulus has been perceived by answering "YES" or "NO"; (- - - -) sensory threshold.
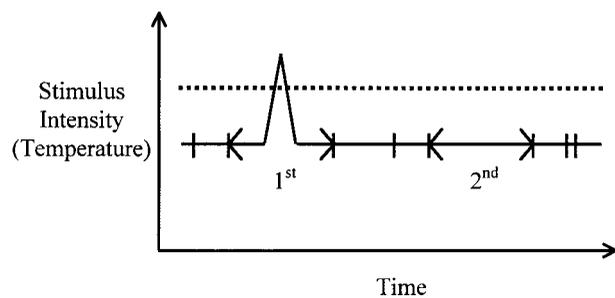
FIGURE 3. Method of levels with temporal forced-choice technique. (|) Warning prior to time interval that may or may not contain the actual stimulus; (|↔|) time interval; (⋀) stimulus of predetermined intensity; (‖) subject is forced to choose which of the two time intervals contain the stimulus; (- - - -) sensory threshold.

Since the answer is not dependent on the reaction time, this algorithm is considered reaction time–exclusive. The only drawback to this method is that the subject has to remember the baseline "internal noise" to determine when something more (a stimulus) has been added.

Forced-Choice Method. The forced-choice method reduces the need for the subject to set or remember the level of "internal noise." It was first introduced into clinical use by Sekular and colleagues in 1973 as a sensitive and objective procedure for evaluating response to light touch.[66] In 1978, Dyck and colleagues introduced the temporal forced-choice paradigm in QST.[32] Using this method, the subject is first presented with two time periods (Fig. 3). Only one of the two time periods contains the stimulus of predetermined intensity (method of levels). The subject is then "forced" to choose which time period contains the stimulus (first or second).

Another forced-choice method is called the spatial forced-choice paradigm. Here, the subject is presented with two stimulating probes with the stimulus present at only one of them. The subject is asked to touch each probe and identify that with the stimulus.[7,11,21,55] Inherent in the forced-choice method is the use of the method of levels and not the method of limits.

The use of null stimuli, or time period with no actual stimulus, may be applied to the algorithm to improve the accuracy of the test. When a subject identifies a stimulus during a time period when no stimulus was given, the ability of the subject to distinguish stimulus from background noise is called into question.

With the forced-choice method, the subject is required to choose which time period (temporal) or stimulating probe (spatial) contains the stimulus even when the stimulus level is below threshold or

when no stimulus is actually present (null stimulus). Some subjects may find this protocol challenging. Another disadvantage of the forced-choice method is that it is time consuming.

It is clear from the above discussion that the forced-choice technique of eliciting subjects' responses may be used in different ways: temporal, spatial, and with or without null stimuli. Simply using the term "forced-choice method" to describe the entire process of eliciting the subject response is confusing and should be discouraged.

How the Subject's Responses Are Utilized to Determine Sensory Threshold.
The following are brief descriptions of the different ways of determining sensory thresholds used by different investigators.

Method of Limits (Thermal). Claus and colleagues[17] started from the reference temperature of 35°C. Seven warm or cold stimuli with a linear ramp rate of 1°C/s were presented. The subject was instructed to press a button as soon as a temperature change was perceived and the thermode temperature returned to the reference temperature. The difference between the temperature achieved and the reference temperature served as a reading. After eliminating the first trial, the average of six warm (or cold) readings was considered the threshold. The warm and cold sensory thresholds were determined separately.

Hilz and colleagues[44] started from a baseline temperature of 32°C. Five warm or cold stimuli with a ramp rate of 1°C/s were presented to the subject. They then averaged the threshold from the responses to five consecutive stimuli. Yarnitsky and Sprecher[80] started from a reference temperature of 32°C and considered the average of four readings to be the sensory threshold.

Fagius and colleagues[33] used the method of Fruhstorfer and colleagues.[36] They considered the warm–cold difference limen or the temperature interval between the perception of warm and cold to be the sensory threshold. They did not determine the warm and cold thresholds separately with this method. This method is also called the original Marstock method (Fig. 1A), as described earlier.

Method of Limits (Vibration). Claus and colleagues[18] asked the subject to press a button as soon as the vibration was perceived during an increasing ramp (perception threshold) and as soon as vibration disappeared with a decreasing ramp (disappearance threshold). They considered the mean of three perception thresholds and three disappearance thresholds to be the vibration perception threshold.

Armstrong and colleagues[8] considered the vibration perception threshold to be the mean of three trials, whereas both Bertelsmann and colleagues[10] and Guy and colleagues[39] considered it to be the mean of five trials.

Fagius and colleagues[33] determined the vibration perception threshold by increasing the vibration intensity from zero until it was felt by the subject ("up" value) and by decreasing the stimulus from a supraliminal intensity until it disappeared ("down" value). They found that the responses usually became stable after two or three repetitions, and the threshold was the average of at least two up and down values.

Method of Levels (Vibration, Thermal). Yarnitsky and Sprecher[80] determined thermal threshold by giving the subject a series of stimuli of varying predetermined intensities. After an initial temperature step of 4°C, the subject was asked to give a "yes" or "no" response to indicate whether the stimulus was perceived. A "yes" response led to a smaller step, and a "no" response led to a larger stimulus. The step size was halved when there was a change in subject's response from "yes" to "no," or vice versa. The step size remained unchanged when the subject's response remained unchanged. The process was continued until the step size reached 0.2°C. Yarnitsky and Sprecher used the average of stimulus temperature for the last "yes" and the last "no" as the threshold.

The staircase algorithm was described by Fowler and colleagues[34] to determine thermal thresholds. First, the subject received thermal stimuli at a level well above the threshold. The next level of stimulus intensity depended on the subject's response: it was reduced by one step level if the subject answered "yes" and increased if the response was "no." The test was terminated when four "no" responses were given. Fowler and colleagues considered the midpoint between the mean of values with "yes" responses following the first "no" and the mean of values with "no" responses as the subject's threshold.[34]

When the stimulus is presented to the subject using the method of levels, Dyck and colleagues[32,59] described two ways of determining vibration and thermal sensory thresholds.

(a) *Temporal forced-choice algorithm.* The up–down transformed response rule as proposed by Wetherill was selected by Dyck and colleagues to find the sensory threshold.[32] This rule was used by Dyck to move the stimulus levels up and down to estimate the sensory threshold. Under the forced-choice paradigm, the sensory threshold corresponds to the stimulus level for which the subject's responses are

correct 75% of the time (correct responses may occur by chance alone 50% of the time when no stimuli are actually present, or when the subject felt the actual stimuli 50% of the time).[59]

Using the forced-choice paradigm as described earlier, the subject was first presented a stimulus of a predetermined level of intensity. The next level of stimulus intensity depended on whether the subject was able to perceive the first stimulus. When the response given was correct, the next level of stimulus intensity was reduced until an incorrect answer was obtained, which led in turn to a higher intensity stimulus. This level of stimulus intensity was called the turnaround value. In order to reduce the chance that the response was not just a correct guess (50% of the time), a certain combination or series of correct or incorrect answers were used to determine the sensory threshold. There are several ways to define success at a given level of stimulus intensity.

Dyck and colleagues developed the CASE (computerized assisted sensory examination) system of quantitative sensory testing.[32] This system has undergone a number of modifications in machine specifications and algorithms. Under the CASE III system, an observed sequence of SSS or SSFS (S = success, F = failure) was considered a correct response and this would lead to a stimulus of lower intensity. An observed sequence of F, SF, SSFF was considered incorrect, leading to a higher stimulus intensity.

In the most recent CASE IV system, the definition of success was changed to improve the accuracy of the estimates.[59] Success is now defined as giving six correct answers out of the first seven trials or five correct answers in the first six trials with the seventh trial incorrect and the eighth and ninth trials correct. Due to the wide range of thermal stimuli, the available levels to be tested were reduced to 25 defined levels of stimulus intensity that were based on just-noticeable differences (JNDs) in healthy subjects. Steps 1 through 21 consist of pyramidal-shaped heat pulses, starting from a baseline temperature of 34°C, in which the temperature increases or decreases at a constant rate of 4°C/s. At step 21, the pyramidal-shaped heat pulse reaches a peak of 48°C. The peak temperature of 48°C is sustained for 1.5 s in step 22, 5 s in step 23, and 10 s in step 24. At step 25, the temperature is sustained for 10 s at 49°C. The test starts at an intermediate nonpainful stimulus intensity (level 13). Subsequent changes in stimulus intensity were originally made in steps of 4, 4, 2, 2, 1, 1, 1, 1, preceding the eight turnaround points, respectively. This sequence was changed in CASE IV so that, after the first failure, subsequent decreases are

in single steps. The threshold is the median of the last six turnaround values.

(b) *One-time-period 4, 2, 1 stepping algorithm with null stimuli.* The one-time-period 4, 2, 1 stepping algorithm with null stimuli algorithm was introduced by Dyck and colleagues in 1993 as a faster way to estimate sensory threshold[30]; compared with the forced-choice method, this algorithm is much less time consuming. The test starts with level 13. Subsequent changes in stimulus intensity depend on whether the stimulus is felt occurring in steps of 4 until the first turnaround, in steps of 2 until the second turnaround, and in steps of 1 thereafter. A total of 15 stimulus events and 5 null stimuli (randomly interspersed) are used. The mean of the turnaround levels with single steps is taken as the threshold. This algorithm is now part of the CASE IV system. A positive response to two null stimuli prompts reinstruction and retesting of the subject. Repeated positive responses from the subject to null stimuli indicate that this algorithm cannot be used for that individual subject.[30] Additional testing of these subjects will require the use of the forced-choice method described earlier.

A spatial forced-choice algorithm was described by Nasseri and colleagues to determine vibration threshold. Using the Vibratron II (Physitemp, Clifton, NJ), the test began at the highest intensity of 20 vibration units (VU) or at a lower intensity level appropriate for the subject. The intensity was lowered by 10% after each correct response and was increased by 10% after each incorrect response. Five incorrect responses were enough to complete the examination. All levels below 1 VU were repeated twice. The lowest and highest values of ten incorrect and correct responses were eliminated and the mean of the remaining eight scores was called the vibration threshold.[58]

Recognizing that there are several different algorithms in use, the American Diabetes Association made a number of recommendations to clinical investigators in 1992 regarding QST as it applies to diabetic neuropathy. The two testing algorithms recommended were: (1) the two-alternative forced choice; and (2) the method of limits combined with yes–no paradigm with null stimuli.[6] Of note, the one-time-period 4, 2, 1 stepping algorithm with null stimuli was introduced later in 1993, so it was not considered.

HEAT-PAIN SENSATION

Heat-pain thresholds may be determined with the method of limits or the method of levels.

Method of Limits. The methods of limits for heat-pain threshold has been described by Verdugo and Ochoa[73] and Yarnitsky and colleagues.[82] Verdugo and Ochoa[73] started from a baseline temperature of 32°C. They decreased the temperature (cold pain) or increased (heat pain) it at a rate of approximately 4°C/s. The subject was instructed to press the switch at the instant that pain was first felt. The average of three to six measurements was considered as the mean pain threshold. Cold-pain hyperalgesia was diagnosed when the mean cold-pain threshold was higher than 21.1°C on the thenar eminence, 25.5°C on the hypothenar eminence, and 23.5°C on the tarsal region. Heat-pain hyperalgesia was diagnosed when the mean heat-pain threshold was below 40.8°C, 37.2°C, and 38.7°C at these three sites, respectively. Cold- and heat-pain hypoalgesia were not considered by Verdugo and Ochoa because several normal subjects had no cold or heat pain.

Yarnitsky and colleagues[82] started from a baseline temperature of 32°C. The temperature increases were at a rate of 2°C/s. The subject was instructed to depress a switch as soon as pain was perceived. They considered the average of three measurements taken at 20-s interstimulus intervals to be the heat-pain threshold.

Method of Levels. The only study using the method of levels was done by Dyck and colleagues. They called the method the nonrepeating ascending algorithm (NRA-NS).[31]

Nonrepeating Ascending Algorithm. The CASE IV system uses the nonrepeating ascending algorithm to determine the heat-pain threshold.[31] Testing for heat-pain detection threshold requires that the number of stimuli are minimized primarily because an excessive number of heat-pain stimuli may cause local tissue changes that in turn may alter the sensory threshold. The use of forced-choice or 4, 2, 1 stepping algorithms both require repeated stimulations, and these were thought to be inappropriate for heat-pain testing.[31]

In the NRA-NS method, the subject is first given a stimulus at a predetermined level of stimulus intensity and duration. Using a modified yes–no response paradigm with a visual analog scale, the subject is asked to grade stimuli from 0 to 10; with 0 being a nonpainful stimulus and 1 being the least painful stimulus. If the subject rates the stimulus as less than 5, subsequent stimuli are of progressively higher intensities, without repeating any stimulus at the same intensity. Once the subject rates the stimulus as 5 or higher, no subsequent stimulus is given.

The heat-pain detection threshold (HPDT), also called HP:0.5, is the midpoint between the stimulus level graded as nonpainful, and the level graded as 1. The HP:5.0 is the stimulus intensity that is felt to be the intermediate pain level. Based on the actual ratings of the subject, the HP:0.5 and HP:5.0 are usually interpolated from a quadratic equation. The difference between HP:5.0 and HP:0.5 may also be calculated. HP:0.5, HP:5.0, and HP:5.0 minus HP:0.5 are expressed in JNDs and in percentile of normal subjects.[31] Testing for cold-pain threshold has not been standardized under the CASE IV system. This is thought to be more variable than heat-pain threshold (Fig. 4).[38]

EQUIPMENT

A summary of the different equipment and their specifications as reported in the literature can be found at www.aaem.net/aaem/practiceissues/technologyreviews/technologyreviews.cfm.

REPRODUCIBILITY

The increasing use of QST may give the impression that it is reliable and easily reproducible. On the contrary, the results of reproducibility studies appear
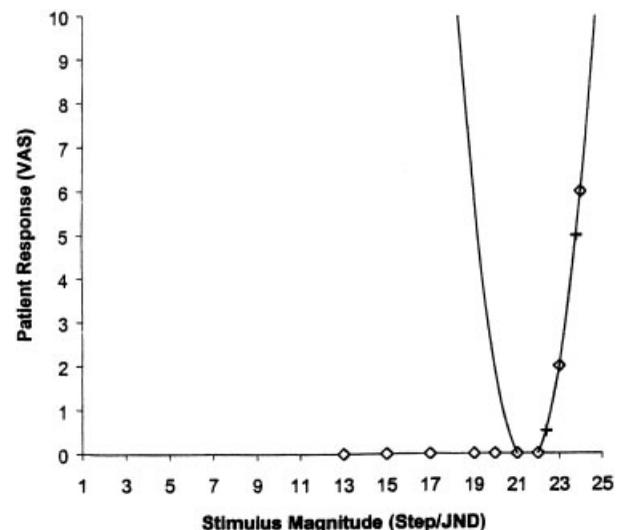
FIGURE 4. Heat-pain threshold (CASE IV). (◇) Patient's response using a visual analog scale (VAS) to grade painful stimuli from 0 to 10, with 0 being a nonpainful stimulus, 1 being the least painful stimulus, and 10 being the most painful stimulus. (+) HP:0.5 and HP:5.0 are interpolated from a quadratic equation (— fitted curve). HP:0.5 is the midpoint between a nonpainful stimulus and the least painful stimulus. HP:0.5 is the estimated heat-pain detection threshold (HPDT) and it corresponds to the stimulus that is felt to be painful 50% of the time. HP:5.0 is the stimulus intensity felt to be the intermediate pain level.

to be confusing, with some studies showing excellent results and others showing poor reproducibility. This problem is compounded by the use of different equipment, algorithms, populations, and statistical methods to measure reproducibility.

QST is not a specific test of peripheral nerve function. Central nervous system dysfunction such as stroke and multiple sclerosis may produce QST abnormalities. The lack of any single "gold standard" for the diagnosis of neuropathy also adds to its nonspecificity. However, QST is strongly correlated with NCSs and with patients' symptoms in diabetic neuropathy. The two important prerequisites of a useful clinical neurophysiologic test are the availability of normal values and the reasonable reproducibility of the test. Several investigators have published normal values for vibration threshold, cold perception threshold, and warm perception threshold. The effect of age, sex, and site of stimulation have been incorporated in the normative data. It should be emphasized that any deviations in equipment, methodology, or algorithm may change the results significantly.

The use of QST to follow patients over time requires that the test is reproducible. We reviewed the literature and summarized the reproducibility studies that have been reported over the past several years. The following factors may influence reproducibility: (1) equipment; (2) thermode size; (3) methodology (method of limits vs. method of levels, yes–no method vs. forced-choice method, "4, 2, 1 stepping algorithm" vs. "up and down transformed response rule"); (4) population studied (age, sex, normal subjects vs. patients, or trained normal subjects vs. untrained normal subjects); (5) number of subjects studied; (6) number of examiners or centers involved (one vs. several); (7) baseline skin temperature; (8) stimulus characteristics (stimulus duration, rate of temperature change, thermode size); (9) stimulus sites; and (10) duration of intervals between tests (hours, days, weeks, or months).

As stated earlier, the 4, 2, 1 stepping algorithm was introduced by Dyck and colleagues to reduce the time needed to complete QST.[30] It should be emphasized that this algorithm was not used in the reproducibility studies on vibration and thermal thresholds.[27,29] Although this new algorithm has been shown to be comparable to the forced-choice algorithm,[30] the reproducibility of the two algorithms may not be identical.

Dyck and colleagues studied the influence of different stimulus characteristics and algorithms of testing on sensory threshold determination. They found the following:

1. The inclusion of reaction time in the algorithm gives a higher sensory threshold compared with algorithms that exclude reaction time.
2. The rate of temperature change of the thermode also affects sensory thresholds: rates of 1°C/s and 3°C/s give higher thresholds than rates below 1°C/s.
3. A larger thermode size gives lower thermal thresholds (10 cm² vs. 2.7 cm²).
4. There is a topographic difference of cold perception threshold, warm perception threshold, and heat-pain detection threshold.
5. Thresholds are lowest in the face and volar arms and highest in the legs and feet.
6. Warm threshold varies among different sites and may be absent in the dorsal foot of older subjects, in whom the first sensation felt is heat pain.[28]
7. The number of examiners involved in the performance of QST may be important.

It should be noted that the study by Jamal and colleagues indicated that only one person did all the QSTs.[48,50] This is rarely possible in a neurophysiology laboratory. Therefore, the reproducibility may be different if several technicians are performing the QST. It is of interest to note that the two multicenter studies[8,71] both showed poor reproducibility of thermal perception thresholds. This may very well be related to the larger number of investigators involved.

The following seven different statistical tools have been used to measure intraindividual reproducibility: (1) correlation coefficient; (2) coefficient of variation; (3) coefficient of repeatability or r-value; (4) intraclass correlation coefficient; (5) ratio of subsequent measurement/initial measurement; (6) percentage of time the measurements are within a certain "number of stimulus steps"; and (7) percent change from the initial measurement.

The choice of statistical method to measure reproducibility is important. What can be applied to a group cannot be applied to an individual patient. Knowing that the mean threshold of a group of individuals changed by less than 5% between two trials[48] may be important for a group study, but it is not helpful for a clinician who is trying to decide whether a specific change in a patient's threshold is clinically significant. Likewise, the correlation coefficient, intraclass correlation coefficient, or coefficient of variation cannot be utilized by clinicians to determine how much change is necessary to signify an improvement or worsening. The use of percent change from initial measurement,[33] 90th percentile absolute day-to-day differences,[18] and of r-values[80]

would be more practical for the clinician. The *r*-value means that there is 95% confidence that two measurements from the same patient differ by less than that value.[80]

A table that summarizes the different reproducibility studies on vibration, thermal, and heat-pain thresholds that have been reported since 1981 can be found at www.aaem.net/aaem/practiceissues/technologyreviews/technologyreviews.cfm. It is evident that the reproducibility of vibration threshold is generally good with both the method of limits[8,10,18,27,39] and the method of levels.[18,27,29]

The first reproducibility study on thermal testing was very disappointing.[33] In the study, the warm–cold difference limen was obtained with the method of limits and used as the thermal threshold.[33] Since then, several studies using the method of levels have shown satisfactory reproducibility.[7,11,14,29,48] The improved reproducibility may be secondary to better algorithms, stimulators, and computers, and the use of the reaction time–exclusive method (method of levels). However, as stated earlier, two multicenter studies showed poor reproducibility despite the use of reaction time–exclusive methods,[8] probably because a larger number of investigators were involved. It is not clear why some studies of thermal threshold using reaction time–exclusive methods showed poor reproducibility.[17,21] Since different machines and methodologies were used, one can only speculate that these factors account for the conflicting results.

A few investigators have compared methods of limits with methods of levels. The reproducibility of thermal and vibration thresholds with both methods were similar in the study by Claus and colleagues.[17,18] Yarnitsky and colleagues[80] found the method of levels to be more reproducible than the method of limits. These apparently conflicting results may be related to the duration of testing. Claus and colleagues found that forced-choice testing at one site took more than 30 minutes to complete in normal subjects and 40 minutes in diabetics.[17] The long duration of testing can lead to inattention and spurious results, thus offsetting whatever advantages forced-choice testing may have over the method of limits.

The site of stimulation also affects the reproducibility of thermal testing. Bravenboer and associates found that the reproducibility of warm and cool thresholds was good in the hand but poor in the foot. Warm threshold in the foot could not be determined in 41% of their diabetic subjects (without neuropathy). This difference may be caused by a more heterogeneous distribution of receptors in the foot or other factors such as skin temperature.[14] In

fact, some normal subjects have also been found to have no warm sensation in the foot. In such subjects, the first sensation in the feet was heat pain rather than warm sensation.[38]

The time interval between tests is also an important factor. Fagius and colleagues found that vibratory thresholds are more reproducible with short intervals.[33] It is important to note that most of the reproducibility studies used test intervals that ranged from a few days to a few weeks, and this may not be applicable to drug trials that may take months to years.

The reproducibility of QST in normal subjects is probably better than that of patients with neuropathy.[10] Moreover, intersession bias may be due to a training or learning effect.[80] Drug trials in patients with neuropathy should include serial QSTs of the placebo group as well as the treatment group.

A consensus report from the Peripheral Neuropathy Association concluded that the stimulus waveform produced by von Frey filaments is too variable and not sufficiently quantifiable to provide the best assessment of touch-pressure sensation.[62] Recently, however, another reproducibility study of the Semmes–Weinstein monofilaments has shown promising results.[72] In a group of 68 diabetics, the coefficient of variation was 0.41 and the correlation coefficient was 0.80 when a group of 68 diabetics were studied by one observer testing on 2 separate days with an interval of 2–4 weeks.[72]

The two reproducibility studies on heat-pain threshold used different machines and methodologies. Both studies revealed results that are sufficiently reproducible[29,79] (see the table on reproducibility studies at the AAEM website: www.aaem.net/aaem/practiceissues/technologyreviews/technologyreviews.cfm).

In summary, vibration perception threshold, heat-pain threshold, cold perception threshold, and warm perception threshold of normal subjects appeared to be sufficiently reproducible during short-term studies (1–8-week intervals). Compared with the vibration perception threshold, thermal thresholds appeared to be more susceptible to different factors such as the machine used, methodology, duration of testing, and time interval between tests. Cool threshold in the foot appears to be more reproducible than warm threshold. Any therapeutic trials with longitudinal QST studies should be performed with a control group of abnormal subjects to demonstrate that the observed changes (improvement or deterioration) in QST results cannot be attributed to the technique's inherent variability. With regard to follow-up studies of the individual patient, the use of *r*-values or repeatability coeffi-

cients[14,80] has practical utility. The recent reproducibility study of the Semmes–Weinstein monofilaments is encouraging but further studies are necessary.

CLINICAL USES OF QST

Diabetic Neuropathy. The painless, noninvasive nature of QST makes it an attractive tool for therapeutic trials and studies on the natural history and progression of diabetic neuropathy.[1,9,43,52,68,69] Both large- and small-fiber functions may be evaluated by QST. Since the introduction of the first Marstock thermal stimulator, we have seen a marked increase in the use of QST in the research and clinical evaluation of patients with diabetic neuropathy. The minimum criterion of diabetic neuropathy as proposed by Dyck and colleagues includes QST as one of the five validated tests.[28] Both the Peripheral Neuropathy Association and the American Diabetes Association included QST in the neurophysiologic evaluation of patients with diabetic neuropathy.[5,6,62]

Many studies have reported on the frequency of QST abnormalities in diabetic neuropathy. Studies on diabetics with minimal or no symptoms have shown a higher frequency of thermal testing abnormalities compared with vibratory or nerve conduction abnormalities.[42,51,53,58,84] This finding suggests an early involvement of the small sensory fibers. Some studies, however, have shown that nerve conduction studies[23,54,64] and vibratory perception threshold[23,26,64,67,74,76] are more frequently abnormal than thermal testing.

These apparently conflicting results are probably related to the characteristics of the population studied. Patients with predominantly small-fiber dysfunction should have a higher frequency of abnormalities in thermal testing,[49,85] whereas those with large-fiber involvement should show abnormalities in vibration threshold.[51,56] In the advanced stage of diabetic neuropathy, both small and large fibers are more uniformly affected and the diagnostic yield of NCSs, and vibratory and thermal testing are all higher.[53] Therefore, there is no single most sensitive electrophysiologic test for diabetic neuropathy. Instead, one should be guided by the clinical characteristics of the diabetic neuropathy.

There are also conflicting results as to which thermal test (warm vs. cool threshold testing) is more likely to be abnormal in diabetic neuropathy. Ziegler and colleagues studied a group of 40 newly diagnosed type 1 diabetic patients using the Marstock method and found the following abnormalities: cool threshold (thenar and foot) in 27.5%;

warm threshold in 22.5% on the thenar surface and 12.5% on the foot; vibration threshold in 7.5% (foot); and thermal discriminating threshold or warm–cold difference limen in 25%.[84] Another study involving 25 patients with small-fiber neuropathy (half of whom were diabetics) showed abnormalities of warm threshold in 88% and of cold threshold in 72%.[49] The following relative frequencies of abnormal thermal tests were noted in a study of 280 diabetics: thermal sensitivity limen or Marstock's warm–cold difference limen, 79.2%; cool threshold, 68%; and warm threshold, 59.2%.[58] Five of 11 diabetics with clinically insignificant neuropathy and normal cool threshold had abnormal warm threshold.[53] The percentages of warm and cold threshold abnormalities were very similar (78% and 77%, respectively) in the study by Vinik and colleagues involving 81 diabetics.[76] This contrasts with another study of 60 diabetics that showed abnormal cool threshold in 98%, and abnormal warm threshold in 58%.[42] Since there is no consensus as to which modality is more sensitive, the choice will depend on the reproducibility of the selected protocol. As mentioned earlier, the reproducibility of warm threshold appears to be poor in the foot.

Studies of patients with painful diabetic neuropathy have shown elevation of both thermal and vibration thresholds.[41,85] Patients with painful neuropathy tend to have more pronounced abnormalities of thermal testing than of vibration[39,41] or NCSs.[85] This is not surprising since thermal testing evaluates small-fiber function, which is usually abnormal in painful neuropathy. In contrast, Veves and colleagues found no significant difference in QST between diabetics with painless and painful neuropathy.[75] The reason for these conflicting results is not known. Differences in population characteristics, machines, and testing protocols may all contribute to the differing results.

Diabetics with foot ulceration have been found to have markedly abnormal cool, warm, and vibration thresholds in the feet.[3,13,45] In one study, 3 of 20 patients had abnormal warm thresholds with minimal cool threshold abnormalities.[3] Impaired pressure sensation as detected by the Semmes–Weinstein pressure aesthesiometer has also been associated with foot ulceration in diabetic patients.[60,70] Diabetic neuropathy is a heterogeneous disorder. Large, small, myelinated, and unmyelinated nerve fibers are affected in different proportions. Additionally, the progression of diabetic neuropathy is probably not only a length-dependent centripetal process,[65] but also a random one,[77] with a higher probability of affecting the longer nerve fibers in the lower extrem-

ities. Rather than using one or two sensory modalities, it is more logical to use a number of quantitative tests to follow the course of the disease. The use of a composite score (NIS[LL]+7 tests), which includes QST and NCSs, has been described recently to assess the severity of diabetic neuropathy longitudinally for controlled clinical trials.[24]

Carpal Tunnel Syndrome. The current "gold standard" for the neurophysiologic diagnosis of carpal tunnel syndrome (CTS) is NCSs. Median sensory, mixed, and motor NCSs can easily demonstrate nerve dysfunction across the wrist. The specificity of NCSs in the diagnosis of CTS is increased significantly when the ulnar nerve is found to be normal, thus ruling out generalized neuropathy.

Since large myelinated fibers are more susceptible to compression, one would expect vibratory threshold testing to be quite sensitive in CTS. Unfortunately, several QST studies have shown that a significant percentage of patients with CTS have elevated vibratory thresholds not only in digits 2 and 3 (median innervated) but also in digit 5 (ulnar innervated).[12,37,47,57] The low sensitivity and specificity of QST in the diagnosis of CTS limits its clinical utility. The use of QST in the diagnosis of CTS has not been fully established.[4]

Other Potential Uses. Aside from documenting sensory deficits (hypoesthesia), there are many other potential clinical uses of QST. QST is the best test available for documenting hyperesthesia.[73] Neuropathy secondary to alcohol abuse, renal failure, human immunodeficiency virus infection, paraneoplastic syndrome, chemotherapy exposure, immune-mediated disorders, hereditary neuropathy, vitamin B_{12} deficiency, toxin exposure, leprosy, and connective tissue disease may be followed longitudinally to detect any worsening or response to therapy. Sensory nerve function may also be followed after nerve trauma and repair. QST abnormalities may also be noted with central nervous system disorders such as multiple sclerosis, syringomyelia, and cerebrovascular lesions. The role of QST in radiculopathy and entrapment neuropathies other than CTS is still unclear. Recently, Zaslansky and Yarnitsky have provided a comprehensive review of the clinical applications of QST.[83]

Limitations of QST. It is important to emphasize that QST tests the integrity of the entire sensory neuraxis and is of no localizing value. Dysfunction of the peripheral nerves or central nervous system may give rise to abnormalities in QST.

One needs to be familiar with all the potential pitfalls of QST. Being a psychophysical test, QST lacks the objectivity of NCSs. Results are subject to changes due to distraction, boredom, mental fatigue, drowsiness, or confusion. All of the aforementioned factors should be carefully monitored because a "routine" QST that includes vibratory, warm, cold, heat-pain, and cold-pain threshold usually takes 1–2 hours using the 4, 2, 1 algorithm. The forced-choice method takes even longer. In our experience, the forced-choice algorithm, as used in the CASE IV system, may take up to 30 minutes for one modality at one site. Boredom and inattention may give rise to spurious results.

It is not uncommon for patients to become bored after 2 hours of testing. When patients are consciously or subconsciously biased toward an abnormal QST result, no psychophysical testing can reliably distinguish these patients from those with organic disease.[25] Just like any other neurophysiologic test, QST results should always be interpreted in light of the patient's clinical picture.

CONCLUSIONS

After a thorough review of the literature available, it is the opinion of the AAEM that:

1. QST is a reliable psychophysical test of large- and small-fiber sensory modalities.
2. QST tests the integrity of the entire sensory axis from receptors to brain. Abnormalities do not localize dysfunction to the central or peripheral nervous system, or any particular location along the peripheral nervous system.
3. QST is highly dependent on the full cooperation of the patient and may be falsely abnormal if the patient is biased toward an abnormal result or is cognitively impaired. No algorithm can reliably distinguish between psychogenic and organic abnormality.
4. QST has been shown to be reasonably reproducible over a period of days or weeks in normal subjects. Since longitudinal QST studies of patients in drug trials are usually done over a period of several months to a few years, reproducibility studies on the placebo-controlled group should be included.
5. The reproducibility of thermal thresholds may not be as good as that of vibration threshold.
6. For individual patients, more studies are needed to determine the maximum allowable difference between two QSTs that can be attributed to experimental error.

7. Different commercially available QST instruments have different specifications (thermode size, stimulus characteristics), testing protocols, algorithms, and normal values. Only QST instruments and their corresponding methodologies that have been shown to be reproducible should be used for research and patient care.

8. The results of QST can only be interpreted properly if machine calibration and testing protocol are strictly followed.

9. The literature does not allow a conclusion to be made regarding whether any QST instrument is better than another.

DISCLAIMER

This report is provided as an educational service of the AAEM. It is based on an assessment of the current scientific and clinical information. It is not intended to include all possible methods of care of a particular clinical problem, or all legitimate criteria for choosing to use a specific procedure. Neither is it intended to exclude any reasonable alternative methodologies. It addresses the use of QST in one area only and its application, if any, to the practice of electrodiagnostic medicine. This statement is not intended to address all possible uses of or issues regarding QST and in no way reflects upon the usefulness of QST in those areas not addressed. The AAEM recognizes that specific patient care decisions are the prerogative of the patient and his/her physician and are based on all of the circumstances involved. This review was not written with the intent that it be used as a basis for reimbursement decisions. No clinical tests or trials were performed by the AAEM; the AAEM's opinion is based solely on a review of the literature.

REFERENCES

1. Abraham RR, Levy DM, Abraham RM. Changes in thermal sensation in diabetic patients after treatment with gangliosides. Diabetes Res 1988;7:129–135.
2. Adriaensen H, Gybels J, Handwerker HO, Van Hees J. Response properties of thin myelinated (A-delta) fibres in human skin nerves. J Neurophysiol 1983;49:111–122.
3. Ali Z, Carroll M, Robertson KP, Fowler CJ. The extent of small fibre sensory neuropathy in diabetics with plantar foot ulceration. J Neurol Neurosurg Psychiatry 1989;52:94–98.
4. American Academy of Neurology, American Association of Electrodiagnostic Medicine, and American Academy of Physical Medicine and Rehabilitation. Practice parameter for electrodiagnostic studies in carpal tunnel syndrome (summary statement). Neurology 1993;43:2404–2405.
5. American Diabetes Association. Proceedings of a consensus development conference on standardized measures in diabetic neuropathy: quantitative sensory testing. Muscle Nerve 1992;15:1155–1157.
6. American Diabetes Association. Proceedings of a consensus development conference on standardized measures in diabetic neuropathy: summary and recommendations. Muscle Nerve 1992;15:1167–1170.
7. Arezzo JC, Schaumburg HH, Laudadio C. Thermal sensitivity tester. Device for quantitative assessment of thermal sense in diabetic neuropathy. Diabetes 1986;35:590–592.
8. Armstrong FM, Bradbury JE, Ellis SH, Owens DR, Rosen I, Sonksen P, Sundkvist G. A study of peripheral diabetic neuropathy. The application of age-related reference values. Diabetes Med 1991;8(suppl):S94–S99.
9. Bertelsmann FW, Heimans JJ, van Rooy JC, Dankmeijer HF, Visser SL, van der Veen EA. Peripheral nerve function in patients with painful diabetic neuropathy treated with continuous subcutaneous insulin infusion. J Neurol Neurosurg Psychiatry 1987;50:1337–1341.
10. Bertelsmann FW, Heimans JJ, van Rooy JC, Heine RJ, van der Veen EA. Reproducibility of vibratory perception thresholds in patients with diabetic neuropathy. Diabetes Res 1986;3:463–466.
11. Bertelsmann FW, Heimans JJ, Weber EJ, van der Veen EA, Schouten JA. Thermal discrimination thresholds in normal subjects and in patients with diabetic neuropathy. J Neurol Neurosurg Psychiatry 1985;48:686–690.
12. Borg K, Lindblom U. Diagnostic value of quantitative sensory testing (QST) in carpal tunnel syndrome. Acta Neurol Scand 1988;78:537–541.
13. Boulton AJ, Hardisty CA, Betts RP, Franks CI, Worth RC, Ward JD, Duckworth T. Dynamic foot pressure and other studies as diagnostic and management aids in diabetic neuropathy. Diabetes Care 1983;6:26–33.
14. Bravenboer B, van Dam PS, Hop J, van den Steenhoven J, Erkelens DW. Thermal threshold testing for the assessment of small fibre dysfunction: normal values and reproducibility. Diabetes Med 1992;9:546–549.
15. Chaudhry V, Cornblath DR, Mellits ED, Avila O, Freimer ML, Glass JD, Reim J, Ronnett GV, Quaskey SA, Kuncl RW. Inter- and intra-examiner reliability of nerve conduction measurements in normal subjects. Ann Neurol 1991;30:841–843.
16. Chaudhry V, Corse AM, Freimer ML, Glass JD, Mellits ED, Kuncl RW, Quaskey SA, Cornblath DR. Inter- and intraexaminer reliability of nerve conduction measurements in patients with diabetic neuropathy. Neurology 1994;44:1459–1462.
17. Claus D, Hilz MJ, Neundorfer B. Thermal discrimination thresholds: a comparison of different methods. Acta Neurol Scand 1990;81:533–540.
18. Claus D, Mustafa C, Vogel W, Herz M, Neundorfer B. Assessment of diabetic neuropathy: definition of norm and discrimination of abnormal nerve function. Muscle Nerve 1993;16:757–768.
19. Darian-Smith I, Johnson KO, Dykes R. Cold fiber population innervating palmar and digital skin of the monkey: responses to cooling pulses. J Neurophysiol 1973;36:325–346.
20. Darian-Smith I, Johnson KO, LaMotte C, Shigenaga Y, Kenins P, Champness P. Warm fibers innervating palmar and digital skin of the monkey: responses to thermal stimuli. J Neurophysiol 1979;42:1297–1315.
21. Doeland HJ, Nauta JJ, van Zandbergen JB, van der Eerden HA, van Diemen NG, Bertelsmann FW, Heimans JJ. The relationship of cold and warm cutaneous sensation to age and gender. Muscle Nerve 1989;12:712–715.

22. Duclaux R, Kenshalo DR. Response characteristics of cutaneous warm receptors in the monkey. J Neurophysiol 1980;43:1–15.
23. Dyck PJ, Bushek W, Spring EM, Karnes JL, Litchy WJ, O'Brien PC, Service FJ. Vibratory and cooling detection thresholds compared with other tests in diagnosing and staging diabetic neuropathy. Diabetes Care 1987;10:432–440.
24. Dyck PJ, Davies JL, Litchy WJ, O'Brien PC. Longitudinal assessment of diabetic polyneuropathy using a composite score in the Rochester Diabetic Neuropathy Study cohort. Neurology 1997;49:229–239.
25. Dyck PJ, Dyck PJ, Kennedy WR, Kesserwani H, Melanson M, Ochoa J, Shy M, Stevens JC, Suarez GA, O'Brien PC. Limitations of quantitative sensory testing when patients are biased toward a bad outcome. Neurology 1998;50:1213.
26. Dyck PJ, Karnes JL, Daube J, O'Brien P, Service FJ. Clinical and neuropathological criteria for the diagnosis and staging of diabetic polyneuropathy. Brain 1985;108:861–880.
27. Dyck PJ, Karnes JL, Gillen DA, O'Brien PC, Zimmerman IR, Johnson DM. Comparison of algorithms of testing for use in automated evaluation of sensation. Neurology 1990;40:1607–1613.
28. Dyck PJ, Karnes JL, O'Brien PC, Zimmerman IR. Detection thresholds of cutaneous sensation in humans. In: Dyck PJ, Thomas PK, Griffin J, Low PA, Poduslo JF, editors. Peripheral neuropathy. Vol. 1. Philadelphia: W. B. Saunders; 1993. pp 706–728.
29. Dyck PJ, Kratz KM, Lehman KA, Karnes JL, Melton LJ, O'Brien PC, Litchy WJ, Windebank AJ, Smith BE, Low PA, Service FJ, Rizza RA, Zimmerman BR. The Rochester diabetic neuropathy study: design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests. Neurology 1991;41:799–807.
30. 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.
31. Dyck PJ, Zimmerman IR, Johnson DM, Gillen D, Hokanson JL, Karnes JL, Gruener G, O'Brien PC. A standard test of heat-pain responses using CASE IV. J Neurol Sci 1996;136:54–63.
32. Dyck PJ, Zimmerman IR, O'Brien PC, Ness A, Caskey PE, Karnes J, Bushek W. Introduction of automated systems to evaluate touch-pressure vibration, and thermal cutaneous sensation in man. Ann Neurol 1978;4:502–510.
33. Fagius J, Wahren LK. Variability of sensory threshold determination in clinical use. J Neurol Sci 1981;51:11–27.
34. Fowler CJ, Carroll MB, Burns D, Howe N, Robinson K. A portable system for measuring cutaneous thresholds for warming and cooling. J Neurol Neurosurg Psychiatry 1987;50:1211–1215.
35. Fowler CJ, Sitzoglou K, Ali Z, Halonen P. The conduction velocities of peripheral nerve fibres conveying sensations of warming and cooling. J Neurol Neurosurg Psychiatry 1988;51:1164–1170.
36. Fruhstorfer H, Lindblom U, Schmidt WG. Method for quantitative estimation of thermal thresholds in patients. J Neurol Neurosurg Psychiatry 1976;39:1071–1075.
37. Goadsby PJ, Burke D. Deficits in the function of small and large afferent fibers in confirmed cases of carpal tunnel syndrome. Muscle Nerve 1994;17:614–622.
38. Gruener G, Dyck PJ. Quantitative sensory testing: methodology, applications and future directions. J Clin Neurophysiol 1994;11:568–583.
39. Guy RJ, Clark CA, Malcolm PN, Watkins PJ. Evaluation of thermal and vibration sensation in diabetic neuropathy. Diabetologia 1985;28:131–137.
40. Hallin RG, Torebjork HE, Wiesenfeld Z. Nociceptors and warm receptors innervated by C fibres in human skin. J Neurol Neurosurg Psychiatry 1982;45:313–319.
41. Heimans JJ, Bertelsmann FW, Van Rooy JC. Large and small nerve fiber function in painful diabetic neuropathy. J Neurol Sci 1986;74:1–9.
42. Hendriksen PH, Oey PL, Wieneke GH, Bravenboer B, van Huffelen AC. Subclinical diabetic polyneuropathy: early detection of involvement of different nerve fibre types. J Neurol Neurosurg Psychiatry 1993;56:509–514.
43. Hillson RM, Hockaday TD, Newton DJ. Hyperglycemia is one correlate of deterioration in vibration sense during the 5 years after diagnosis of type 2 (noninsulin-dependent) diabetes. Diabetologia 1984;26:122–126.
44. Hilz MJ, Axelrod FB, Hermann K, Haertl U, Duetsch M, Neundörfer B. Normative values of vibratory perception in 530 children, juveniles and adults aged 3–79 years. J Neurol Sci 1998;159:219–225.
45. Hoeldtke RD, Davis KM, Hshieh PB, Gaspar SR, Dworkin GE. Are there two types of diabetic foot ulcers? J Diabet Complications 1994;8:117–125.
46. Iggo A. Cutaneous thermoreceptors in primates and subprimates. J Physiol (Lond) 1969;200:403–430.
47. Imai T, Matsumoto H, Minami R. Asymptomatic ulnar neuropathy in carpal tunnel syndrome. Arch Phys Med Rehabil 1990;71:992–994.
48. Jamal GA, Hansen S, Weir AI, Ballantyne JP. An improved automated method for the measurement of thermal thresholds. I. Normal subjects. J Neurol Neurosurg Psychiatry 1985;48:354–360.
49. Jamal GA, Hansen S, Weir AI, Ballantyne JP. The neurophysiologic investigation of small fiber neuropathies. Muscle Nerve 1987;10:537–545.
50. Jamal GA, Weir AI, Hansen S, Ballantyne JP. An improved automated method for the measurement of thermal thresholds. II. patients with peripheral neuropathy. J Neurol Neurosurg Psychiatry 1985;48:361–366.
51. Jensen TS, Bach FW, Kastrup J, Dejgaard A, Brennum J. Vibratory and thermal thresholds in diabetics with and without clinical neuropathy. Acta Neurol Scand 1991;84:326–333.
52. Keen H, Payan J, Allawi J, Walker J, Jamal GA, Weir AI, Henderson LM, Bissessar EA, Watkins PJ, Sampson M, and colleagues. Treatment of diabetic neuropathy with gamma-linolenic acid. Diabetes Care 1993;16:8–15.
53. Le Quesne PM, Fowler CJ, Parkhouse N. Peripheral neuropathy profile in various groups of diabetics. J Neurol Neurosurg Psychiatry 1990;53:558–563.
54. Levy DM, Abraham RR, Abraham RM. Small- and large-fiber involvement is early diabetic neuropathy: a study with the medial plantar response and sensory thresholds. Diabetes Care 1987;10:441–447.
55. Levy D, Abraham R, Reid G. A comparison of two methods for measuring thermal thresholds in diabetic neuropathy. J Neurol Neurosurg Psychiatry 1989;52:1072–1077.
56. Maser RE, Nielsen VK, Bass EB, Manjoo Q, Dorman JS, Kelsey SF, Becker DJ, Orchard TJ. Measuring diabetic neuropathy: assessment and comparison of clinical examination and quantitative sensory testing. Diabetes Care 1989;12:270–275.
57. Merchut MP, Gruener G. Quantitative sensory threshold testing in patients with multiple sclerosis. Electromyogr Clin Neurophysiol 1993;30:119–124.
58. Navarro X, Kennedy WR. Evaluation of thermal and pain sensitivity in type I diabetic patients. J Neurol Neurosurg Psychiatry 1991;54:60–64.
59. O'Brien PC, Dyck PJ, Kosenke JL. A computer evaluation of quantitative algorithms for measuring detection thresholds of cutaneous sensation. In: Munsat TL, editor. Quantification of neurologic deficits. Boston: Butterworths; 1989. pp 197–206.
60. Olmos PR, Cataland S, O'Dorisio TM, Casey CA, Smead WL, Simon SR. The Semmes–Weinstein monofilament as a potential predictor of foot ulceration in patients with noninsulin-dependent diabetes. Am J Med Sci 1995;309:76–82.
61. Olney RK, Schleimer JA. Intraexaminer and interexaminer reproducibility of nerve conduction measurements in diabetics with mild polyneuropathy. Muscle Nerve 1992;15:1195.

62. Peripheral Neuropathy Association. Quantitative sensory testing: a consensus report from the Peripheral Neuropathy Association. Neurology 1993;43:1050–1052.

63. Perl ER. Myelinated afferent fibres innervating the primate skin and their response to noxious stimuli. J Physiol (Lond) 1968;197:593–615.

64. Redmond JM, McKenna MJ, Feingold M, Ahmad BK. Sensory testing versus nerve conduction velocity in diabetic polyneuropathy. Muscle Nerve 1992;15:1334–1339.

65. Said G, Slama G, Selva J. Progressive centripetal degeneration of axons in small fibre diabetic polyneuropathy. A clinical and pathological study. Brain 1983;106:791–807.

66. Sekuler R, Nash D, Armstrong R. Sensitive, objective procedure for evaluating response to light touch. Neurology 1973; 23:1282–1291.

67. Sosenko JM, Gadia MT, Natori N, Ayyar DR, Ramos LB, Skyler JS. Neurofunctional testing for the detection of diabetic peripheral neuropathy. Arch Intern Med 1987;147:1741–1744.

68. Sosenko JM, Kato M, Gadia MT, Ayyar DR. Specific assessments of warm and cool sensitivities in adult diabetic patients. Diabetes Care 1988;11:481–483.

69. Sosenko JM, Kato M, Soto R, Bild DE. A prospective study of sensory function in patients with type 2 diabetes. Diabetes Med 1993;10:110–114.

70. Sosenko JM, Kato M, Soto R, Bild DE. Comparison of quantitative sensory-threshold measures for their association with foot ulceration in diabetic patients. Diabetes Care 1990;13: 1057–1061.

71. Valensi P, Attali JR, Gagant S. French Group for Research and Study of Diabetic Neuropathy. Reproducibility of parameters for assessment of diabetic neuropathy. Diabetes Med 1993;10: 933–939.

72. Valk GD, de Sonnaville JJ, van Houtum WH, Heine RJ, van Eijk JT, Bouter LM, Bertelsmann FW. 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.

73. Verdugo R, Ochoa JL. Quantitative somatosensory thermotest. A key method for functional evaluation of small calibre afferent channels. Brain 1992;115:893–913.

74. Veves A, Malik RA, Lye RH, Masson EA, Sharma AK, Schady W, Boulton AJ. The relationship between sural nerve morphometric findings and measures of peripheral nerve function in mild diabetic neuropathy. Diabetes Med 1991;8:917–921.

75. Veves A, Young MJ, Manes C, Boulton AJ. Differences in peripheral and autonomic nerve function measurements in painful and painless neuropathy: a clinical study. Diabetes Care 1994;17:1200–1202.

76. Vinik AI, Suwanwalaikorn S, Stansberry KB, Holland MT, McNitt PM, Colen LE. Quantitative measurement of cutaneous perception in diabetic neuropathy. Muscle Nerve 1995; 18:574–584.

77. Waxman SG. Pathophysiology of nerve conduction: relation to diabetic neuropathy. Ann Intern Med 1980;92:297–301.

78. Williams G, Gill JS, Aber V, Mather HM. Variability in vibration perception threshold among sites: a potential source of error in biothesiometry. BMJ 1988;296:233–235.

79. Yarnitsky D, Fowler CJ. Quantitative sensory testing. In: Binnie CD, Cooper R, Fowler CJ, Mauguiere F, Prior PF, Osselton JW, editors. Clinical neurophysiology—EMG, nerve conduction and evoked potentials. Oxford: Butterworth-Heinemann; 1995. pp 253–270.

80. Yarnitsky D, Sprecher E. Thermal testing: normative data and repeatability for various test algorithms. J Neurol Sci 1994; 125:39–45.

81. Yarnitsky D, Sprecher E, Tamir A, Zaslansky R, Hemli JA. Variance of sensory threshold measurements: discrimination of feigners from trustworthy performers. J Neurol Sci 1994; 125:186–189.

82. Yarnitsky D, Sprecher E, Zaslansky R, Hemli JA. Heat pain thresholds. normative data and repeatability. Pain 1995;60: 329–332.

83. Zaslansky R, Yarnitsky D. Clinical applications of quantitative sensory testing (QST). J Neurol Sci 1998;153:215–238.

84. Ziegler D, Mayer P, Gries FA. Evaluation of thermal, pain and vibration sensation thresholds in newly diagnosed type-1 diabetic patients. J Neurol Neurosurg Psychiatry 1988;51:1420–1424.

85. Ziegler D, Mayer P, Wiefels K, Gries FA. Assessment of small and large fiber function in long term type-1 (insulin dependent) diabetic patients with and without painful neuropathy. Pain 1988;34:1–10.