



ELSEVIER

Diabetes Research and Clinical Practice 54 (2001) 115–128

DIABETES RESEARCH
AND
CLINICAL PRACTICE

www.elsevier.com/locate/diabres

Evaluation of three screening tests and a risk assessment model for diagnosing peripheral neuropathy in the diabetes clinic

David Olaleye ^{a,*}, Bruce A. Perkins ^b, Vera Bril ^c

^a *iBiomatics LLC (a SAS Company), 11000 Weston Parkway, Suite 100, Cary, NC 27513, USA*

^b *Beth Israel Deaconess Medical Center, Joslin Diabetes Center, Harvard Medical School, Boston, MA, USA*

^c *Division of Neurology, EN11-209, Toronto General Hospital/University Health Network, University of Toronto, Toronto, Ont., Canada*

Received 18 September 2000; received in revised form 2 May 2001; accepted 14 May 2001

Abstract

Objective: with the aim of evaluating predictive power, three simple screening tests as alternates to nerve conduction tests for diagnosing diabetic peripheral neuropathy (DPN) were investigated. Results of the screening tests, along with the subjects' demographic and clinical characteristics, were planned as the variables for the development of a risk assessment tool for predicting DPN. **Design:** this is a cross-sectional multi-group comparison study. The study utilized a predictive model derived from one subset of the study population, and prospectively tested in the other subset to predict the presence of neuropathy. **Setting:** Diabetic Neuropathy Research Clinic of the Toronto General Hospital and University Health Network in Toronto, Ontario, Canada from June 1998 to August 1999. **Sample population:** data come from 478 subjects consisting of non-diabetic reference subjects, and patients with type 1 and type 2 diabetes mellitus. **Outcome measures:** nerve conduction studies (NCS) comprised the primary defined outcome. The three screening sensory tests examined in the study were the Semmes–Weinstein 10 g monofilament examination (SWME), superficial pain sensation, and vibration by the *on–off* method. **Results:** the three screening tests are significantly and positively correlated with NCS. An increase in the number of insensate responses in the screening test is associated with an increase in the abnormal NCS score. The strength of the association between NCS and each sensory test was greater when the neuropathy severity stage of the subject was added to the model. Both the SWME and vibration by the *on–off* method tests demonstrated sufficient statistical power to differentiate non-diabetic control subjects from subjects with diabetes, as well as to differentiate subjects with diabetes with and without neuropathy. These two tests, when compared with NCS, also demonstrated acceptable diagnostic performance characteristics in terms of high sensitivity and specificity, total number of correctly predicted cases, and receiver-operating characteristic curves. **Conclusion:** this data, through the development of a model involving training and validation sets, demonstrates that the knowledge of clinical risk factors alters the interpretation of sensory tests for DPN. This finding lends further support to the validity of simple sensory testing maneuvers in the conditional diagnosis of DPN. We recommend annual screening with either the SWME or vibration by the *on–off* method in the primary care and diabetes clinics. © 2001 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Screening for neuropathy; Diagnostic test; Diabetes mellitus; Risk assessment model

* Corresponding author. Tel.: +1-919-653-2730.

E-mail address: david.olaleve@ibiomatrics.com (D. Olaleye).

1. Introduction

Nerve conduction studies (NCS) are the *prima facie* and *de facto* measures used for the diagnosis of diabetic peripheral neuropathy (DPN). The tests are deemed to be objective, reliable, and sensitive, and can be used as statistical instruments or surrogate endpoints for neuropathy in large clinical investigations of DPN [1–5]. NCS have also been recommended in the medical literature as the ‘gold standard’ to evaluate and validate other screening maneuvers that are used to diagnose peripheral neuropathy [5,6]. However, wide use and application of NCS remains limited due to the costs of performing the tests, limited availability of special laboratories equipped to do the tests, and the long waits and inconveniences that the patients undergo to have such tests. The evaluation of peripheral neuropathy in the diabetes clinic is thus a challenging task for physicians.

Clinical practice guidelines recommend annual screening for neuropathy in all patients with diabetes as part of the routine assessment for complications [1,7–9]. Peripheral neuropathy affects 37% of people aged 18 years and above with type 1 diabetes, and at least 20–40% of all those with either Type 1 or 2 diabetes. The diabetes control and complications trial (DCCT) showed that intensive insulin therapy in patients with type 1 diabetes mellitus reduced the risk of developing clinical and electrophysiological deficits of DPN by 61% [10,11]. Early screening and identification of the neuropathic process offer the patient the opportunity to actively alter the course of sub-optimal glycemic control and to take appropriate actions that may lead to significantly reduced morbidity. If neuropathy is detected early, medical surveillance and lifestyle changes decrease the likelihood of foot trauma and other clinical symptoms associated with neuropathy [12–14].

Simple screening maneuvers that are rapid but provide accurate and reliable test results when compared with NCS offer a chance to increase the detection of peripheral neuropathy. Evidence from clinical trials has shown the efficacy of

some strategies to reduce outcomes of amputation and ulceration [12–14]. However, the scope of evidence currently available to discriminate among possible screening methods in order to find an optimal screening test is limited. This may be due in part to study cohorts with restricted spectrum of disease, lack of comparability among screening methods, lack of independently blinded examiners for individual screening tests, and lack of a predictive model with acceptable diagnostic operating characteristics to permit clinical interpretation and applicability of the results [15].

DPN is a major complication of diabetes with social, economic and quality of life cost consequences. By definition, peripheral neuropathy must be detectable by clinical history and physical examination and confirmed by objective diagnostic tests [16]. The primary benefit of a diagnostic sensory test is that it can be used to determine (within a quantifiable margin of error) those patients with DPN in a suspect population prior to initiating a more costly and time consuming procedure such as the NCS. Thus one of the goals of a cross-sectional cohort study [15] was to develop a DPN risk assessment model, reported here, based on each of the three different sensory maneuvers evaluated and to define a diagnostic test with sufficient operating characteristics. While it is clear that the optimal screening test cut-off point, termed the operating point, can ultimately be determined only by means of cost-effectiveness modeling, an alternative approach which achieves a sensitivity and accuracy of greater than or equal to 85% while optimizing specificity can provide an acceptable operating point. Only cost-effectiveness studies that account for both the direct and indirect costs of this diabetes complication can confidently determine whether findings from the screening sensory tests are sufficient to establish a diagnosis of peripheral neuropathy rather than initiating NCS tests after clinical examination.

We evaluated the predictive ability of three simple screening maneuvers, the SWME, superficial pain sensation, and vibration testing by the *on-off* method, using NCS as the gold standard

for DPN in a patient population consisting of non-diabetic reference subjects, and patients with type 1 and 2 diabetes mellitus. The screening tests and clinical characteristics of patients in a training set (also synonymously termed a 'derivation set') were used to develop a risk assessment tool for predicting which patients are more likely to have a diagnosis of DPN. A validation set (also synonymously termed a 'test set') was used to evaluate the operating characteristics of this risk assessment tool. The predictive ability of each sensory test was evaluated using diagnostic performance characteristics such as sensitivity, specificity, total accuracy, likelihood ratios, area under the curve (AUC) and receiver operating characteristics (ROC) curves.

2. Methods

2.1. Patient population

The study was conducted at the Diabetic Neuropathy Research Clinic of the Toronto General Hospital/University Health Network (TGH/UHN) from June 1998 to August 1999. Four hundred seventy eight subjects who met the study eligibility criteria and gave informed consent were used in the study. Approval from the UHN Research Ethics Board was obtained prior to the initiation of the study. A detailed description of the study eligibility criteria and conduct have been published elsewhere [17].

Study subjects provided informed consent and underwent a thorough medical and neurological evaluation in order to exclude neuropathy of other etiologies [17]. Data on standardized bilateral nerve conduction studies, that included motor and sensory nerve tests performed by technologists blinded to the status of the subject, were collected on each patient. The SWME, superficial pain sensation and vibration by the *on-off* method screening tests were performed on each patient by different independent observers blinded to all other test results including physical examination and medical history. Prior to the initiation of the trial, training and practice sessions were held for all examiners to minimize bias in test

reporting and ensure standardization of testing methods for each screening test. The examiners included physicians with different backgrounds (specialists, family physicians), technicians and trained medical assistants.

Of 478 patients, 11% ($n = 52$) had no diabetes (the reference group), 15% ($n = 70$) had type 1 diabetes mellitus, and 74% ($n = 356$) had type 2 diabetes mellitus. Complete data were required for all subjects included in the multivariate analyses.

2.2. Study design

The sample population was partitioned into two groups in order to derive and validate models used to develop the risk assessment tool. Data from the first consecutive 346 patients, the training set, were used to build a risk assessment model designed to help guide physicians in classifying patients with abnormal from normal NCS findings based on the sensory test outcomes and clinical attributes of the patients. Data obtained from the following consecutive 132 patients, the validation set, were used to assess the validity of the classification model constructed with the training data set. Table 1 gives the summary description of the patients contained in both samples.

2.3. Clinical stratification assessment

Based on the neurological history and physical examination performed on the subjects, results showed marked differences in the subject population with respect to stage of neuropathy. A clinical stratification method was employed to categorize the subjects according to neuropathy severity grade. The criteria consisted of signs and symptoms that are indicative of neuropathy.

The neuropathy severity grading was constructed as follows: six symptom scores comprising pain, loss of balance, numbness, tingling, upper limb symptoms, and weakness, with each symptom scored as 1 (present) or 0 (absent); eight reflex scores for bilateral knee and ankle reflexes with each reflex graded as 2 (absent), 1 (reduced) or 0 (normal); five sensory examination signs

Table 1
Patient characteristics by sample classification

Characteristic	Training set	Validation set
Dates of enrollment	22 June 1998–31 March 1999	1 April–30 July 1999
Number of patients	346	132
Nerve conduction studies (NCS) ^a	12.31 ± 8.23 (346)	11.86 ± 8.71 (130)
<i>Screening sensory test^a</i>		
10 g Monofilament	3.53 ± 2.68 (345)	2.78 ± 2.86 (129)
Vibration 1 (by <i>on-off</i> method)	2.32 ± 2.88 (345)	1.96 ± 2.91 (130)
Superficial pain sensation	2.53 ± 2.85 (344)	2.01 ± 2.61 (130)
<i>Neuropathy severity stage (%)^b</i>		
Reference set	9.8 (34)	13.6 (18)
Non-neuropathic	13.3 (46)	26.5 (35)
Mild	20.5 (71)	17.4 (23)
Moderate	24.0 (83)	19.7 (26)
Severe	32.4 (112)	22.7 (30)
<i>Diabetes mellitus status</i>		
Type 1	13.6 (47)	17.4(23)
Type 2	76.3 (264)	69.7 (92)
Reference set	10.1 (35)	12.9 (17)
Age range (year)	55.01 ± 11.70	52.92 ± 11.40
Height (cm)	1.71 ± 0.09	1.69 ± 0.09
HbA _{1c}	7.98 ± 1.67	8.28 ± 1.59
Duration of diabetes mellitus (year)	11.20 ± 11.19	12.53 ± 11.47
<i>Gender (%)^c</i>		
Male	66.8 (231)	63.6 (84)
Female	33.2 (115)	36.4 (48)

Screening tests are population means of number of insensate responses per patient ranging from 0 to 8 for each test. NCS refers to number of abnormal parameters on nerve conduction testing with a maximum possible value of 28 per patient. Numbers in parentheses refer to number of observations in each subgroup. Where range is indicated, values are reported in mean ± S.D. Two subjects in validation set omitted from further analyses due to incomplete data.

^a Two-tailed Student two-sample *T*-test: NCS ($P = 0.60$); monofilament ($P = 0.01$); vibration 1 ($P = 0.23$); superficial pain ($P = 0.07$); age ($P = 0.08$); height ($P = 0.19$); HbA_{1c} ($P = 0.08$); duration of diabetes ($P = 0.25$).

^b χ^2 -test square test: neuropathy stage ($P = 0.004$); diabetes status ($P = 0.334$).

^c Fisher's exact test: gender ($P = 0.52$).

comprising pinprick, light touch, vibration, position sense and temperature, with each sign scored as 1 (present) or 0 (absent). The sensory tests were performed at the toes. All the symptoms and signs were combined for a total of 19 possible points. For the purposes of the baseline analysis, subjects were classified as having 'no neuropathy' if the combined score was between 0 and 5, 6 and 8 for 'mild neuropathy'; 9, 10 and 11 for 'moderate neuropathy'; and 12 or more for 'severe neuropathy'.

2.4. Nerve conduction studies

NCS were performed on each patient using temperature control and fixed distances. Details of the NCS are provided in an earlier publication [17]. Measurements of latencies, distances and amplitudes were done in a standardized fashion using onset latencies, and baseline-to-peak amplitudes. Initial positive peak (if present) to negative peak measurements were conducted for sensory responses. F-waves were generated for all motor

nerves. Conduction velocities were calculated for both motor and sensory nerves.

All conduction velocity and distal amplitude values for the NCS were score 0 for normal and 1 for abnormal. The normal range used for NCS was the mean reference values ± 2 standard deviations (S.D.). The total NCS score for a patient if all the parameters were scored abnormal was 28 points comprising 16 motor and 12 sensory nerve scores. The sum of the abnormal scores was used to define the NCS score. A score of ≥ 3 abnormal parameters in the NCS score defined DPN.

2.5. Sensory testing methods

For each testing modality, the patient was given a reference sensation by application of the stimulus to the sternum. The subject was asked with the eyes closed to describe the sensations experienced sequentially at the sites described [17].

The SWME testing was conducted using a 5.07 size (10 g-force) monofilament applied to a non-callused site on the dorsum of the first toe just proximal to the nail bed and repeated four times on both feet in an arrhythmic manner by an independent examiner. The examiner circled correct responses given by the patient on the scoring sheet and then added the number of errors. The SWME score is defined as the total number of times the application of the monofilament was not perceived by the subject and it varies from 0 to 8.

Superficial pain sensation testing was done by using a sterile Neurotip (Owe Mumford, Oxford, England) applied four times to the 2 sites described for the SWME in an arrhythmic manner. The superficial pain threshold is defined as the total number of times the application of the pain sensation was not perceived and it varies from 0 to 8.

Vibration testing by the *on-off* method was conducted using a 128 Hz tuning fork applied to the bony prominence situated at the dorsum of the first toe bilaterally just proximal to the nail bed. The patient was asked to report the perception of both the start of the vibration sensation and cessation of vibration on dampening. The testing was done twice on each toe and correct responses were recorded on a standardized sheet.

The vibration score is defined as the total number of times the application of the vibrating tuning fork and the dampening of vibration was not felt. The range varies between 0 and 8.

2.6. Statistical analysis

For baseline comparison of the two study cohorts, we used the non-parametric Wilcoxon rank-sum two-sample test for continuous variables, the Pearson's chi-square test (χ^2 -test) for categorical variables, and Fisher's exact test for binary variables. Correlation analysis was performed to examine the relationship between each screening test and NCS. We employed the cluster analytic technique to identify and define patient clusters based on how similar the patients are in terms of their clinical and demographic attributes, position on the sensory test continuum, and NCS findings.

Analysis was performed separately for the training and validation data sets to examine if the two samples reflect different cohorts of patients, and the degree of similarity with respect to patient attributes. All analyses were done on an intention-to-treat basis. Where relevant a *P*-value of 0.05 or less was considered as statistically significant. Data management and analysis were carried out using SAS software version 6.12 [18].

2.7. Construction and validation of the DPN risk assessment model

Data obtained from 346 patients in the training set were used to construct the initial model, while data obtained from the 132 patients in the validation set were used to validate the model. Two patients in the validation set had incomplete data and were excluded from further analysis. Logistic regression was used to develop the prediction model for DPN with the NCS variable (abnormal versus normal) as the dependent variable and the sensory test as the predictor variable. Separate models were constructed for each of the three screening modalities to permit comparison and assessment of the models with regards to which one best predicted the NCS variable.

The logistic regression model was further expanded to include the subject's clinical and demographic characteristics as predictor variables. The variables included in the model were clinical neuropathy stage, gender, age, height, hemoglobin HbA_{1c}, and duration of diabetes. Selection of variables was based on the cluster analysis findings and evidence from the medical literature. Findings from the DCCT study suggest that subjects with clinical peripheral neuropathy were more likely to be older patients, male, taller, with longer diabetes duration, and higher prevalence of retinopathy [19].

For classification purposes, the model that would isolate the largest number of cases with an abnormal NCS finding was used for prediction. If the model predicted a case to have greater than or equal to 75% chance of having an abnormal NCS score then this case was classified as an abnormal NCS finding. After the construction and validation of the initial DPN risk assessment model, the data obtained from all the 478 patients were combined to develop a more extensive, revised prediction model. The model serve as a tool for predicting which patients exhibit higher propensity of having DPN based on the defined clinical characteristics and sensory test scores of subjects. Assessment and evaluation of each screening sensory test's ability to predict NCS was done using the diagnostic performance measures. The measures used are sensitivity, specificity, false positive rate, positive and negative likelihood ratios, overall accuracy, area under the curve (AUC), and receiver operating characteristic (ROC) curves.

A true-positive (TP) finding was defined as an 'abnormal' NCS finding correctly classified with the model; a true-negative (TN) finding was defined as a 'normal' NCS finding correctly classified with the model. A false-positive (FP) finding was defined as a 'normal' NCS finding incorrectly classified by the model, while a false negative (FN) finding was defined as an 'abnormal' NCS finding incorrectly classified. The formulae for computing the rates are sensitivity = $TP/(TP + FN)$, specificity = $TN/(TN + FP)$, and overall accuracy = $(TP + TN)/(TP + TN + FP + FN)$. ROC curves were obtained by varying the stringency of the diagnostic criteria

for each sensory test. The cutoff criterion is defined as the number of insensate responses to the sensory test given by the patient. Thus for the ROC curves, we examined different cutoff points for each sensory test and this was used in the logistic model for predicting whether a patient has an abnormal NCS (that is, ≥ 3 abnormal parameters in NCS score) as opposed to normal NCS finding. AUC was computed by the method of Hanley and McNeil [20].

3. Results

3.1. Patient profile and clinical assessment

The baseline clinical and demographic characteristics of the study patients in the trained and validation samples are shown in Table 1. With the exception of the SWME test and neuropathy severity status variables, the distribution of patients in the two study cohorts did not differ significantly in all other attributes. The mean NCS score for patients in the trained study cohort was 3.7% greater than the NCS score for their counterparts in the validation sample (12.31 vs. 11.86). On average, patients in the trained sample gave 3.5 insensate responses to the SWME questions compared with 2.8 for patients in the validation sample ($P = 0.01$). Many patients in both study cohorts showed various clinical signs or symptoms of peripheral neuropathy with more cases in the 'severe neuropathy' category. A smaller proportion of patients with diabetes was diagnosed as non-neuropathic in the trained data set, 13.3% (46 of 346), compared with 26.5% (35 of 132) of patients diagnosed in the validation sample.

3.2. Correlation between nerve conduction studies and sensory tests

For the total sample, all three screening sensory tests were positively inter-correlated ($P < 0.001$) and also significantly and positively correlated with NCS ($P < 0.001$) (data not shown). The sensory tests' correlation coefficients with NCS are 0.56, 0.54, and 0.57 for the SWME, superficial

Table 2

Diagnostic performance characteristics for Semmes–Weinstein 10 g monofilament sensory test

Model	Sensory test cut off ^a							
	Sensory test only				Sensory test + risk factors			
	2	3	4	5	2	3	4	5
<i>Training model (n = 346)</i>								
Parameter estimate ^b	1.892 ^c	1.822 ^c	2.660 ^c	2.691 ^c	0.915 ^a	1.074 ^a	1.012	0.855
Sensitivity (%)	72	64	40	31	91	89	88	87
Specificity (%)	71	77	95	97	79	79	79	79
Accuracy (%)	72	67	51	44	88	87	86	85
Positive likelihood ratio (%)	2.49	2.83	8.88	10.4	4.28	4.19	4.13	4.09
Negative likelihood ratio (%)	0.40	0.46	0.62	0.71	0.12	0.14	0.16	0.17
AUC (%)	71.5	70.8	67.9	64.2	84.8	83.9	83.1	82.8
<i>Validated model (n = 130)</i>								
Sensitivity (%)	62	58	35	30	82	83	85	87
Specificity (%)	84	92	97	97	76	73	70	70
Accuracy (%)	68	68	53	49	80	80	81	82
Positive likelihood ratio (%)	3.85	7.16	13.1	11.1	3.36	3.06	2.86	2.93
Negative likelihood ratio (%)	0.45	0.46	0.66	0.72	0.24	0.24	0.21	0.18
AUC (%)	73.1	75.0	66.4	63.7	78.7	77.9	77.6	78.7
<i>Revised model (n = 473)</i>								
Parameter estimate ^a	1.963 ^c	2.076 ^c	2.766 ^c	2.709 ^c	0.931 ^b	1.147 ^b	1.257 ^a	1.058
Sensitivity (%)	70	63	39	31	86	86	84	85
Specificity (%)	75	82	96	97	81	82	79	79
Accuracy (%)	71	67	52	45	85	85	83	84
Positive likelihood ratio (%)	2.85	3.57	10.0	10.6	4.45	4.64	3.94	3.99
Negative likelihood ratio (%)	0.40	0.45	0.63	0.71	0.17	0.18	0.20	0.19
AUC (%)	72.7	72.7	67.7	64.2	83.5	83.5	81.4	81.9

^a Sensory test cutoff refers to the number of correct insensate responses. Screening tests are population means of number of insensate responses per patient ranging from 0 to 8 for each test. NCS refers to number of abnormal parameters on nerve conduction testing with a maximum possible value of 28 per patient.

^b Regression model parameter estimate of sensory test on NCS. (a) $P < 0.05$; (b) $P < 0.01$; (c) $P < 0.001$. Risk factors included in the model: neuropathy severity stage, family history of diabetes mellitus, duration of diabetes, age, height, sex, and HbA_{1c}.

pain sensation, and vibration by the *on-off* method, respectively. When all other factors are considered, an increase in the number of insensate responses to the sensory tests is associated with an increase in the NCS score (a score ≥ 3 in the latter indicating abnormal results). The direction and strength of the association between NCS and each of the sensory tests remained unchanged when the study cohort was added to the model. Examining the association between the screening tests and NCS after adding in patient neuropathy severity stage provided further information. For the reference subject population, the association between any sensory test and NCS was not signifi-

cant. On the other hand, the correlation coefficients were positive and increased in magnitude (true for all sensory tests) as the baseline neuropathy condition of the patient worsened. The absolute values of these correlation coefficients range from 0.32 to 0.50.

3.3. Initial model construction and validation

3.3.1. SWME

The first panel in Table 2 displays the diagnostic performance characteristics for the SWME sensory test model using the trained sample data. We compared the test's ability to predict NCS

alone and the effects when other variables were added to the model. The magnitude and sign of the regression estimates in row 1 of panel 1 of the table indicate a strong and positive relationship between the sensory test and NCS. Regardless of the cutoff threshold used, abnormal findings on the SWME test are strong indicators of abnormal NCS findings. After inclusion of other risk factors in the model, at least two or three incorrect responses are necessary in order that the test be able to predict NCS with statistically significant results. The performance characteristics for the test given two insensate responses and without the risk factors in the model are a sensitivity of 72%, a specificity of 71%, and an accuracy of 72% and AUC of 71.5%. When the risk factors are included in the model, there was a statistically significant improvement in the test's ability to predict NCS. For a cutoff threshold of two, both the sensitivity and specificity increased by a magnitude of 19 and 7% points, respectively. The observed percentage points increase in accuracy with AUC parameters of 16 and 13, respectively. The findings are very similar, when we also considered the performance characteristics of the model given a cutoff threshold of three insensate responses.

Panel 2 of the same table shows the results of the validated model for the SWME test. The purpose of the validated model is to validate the results obtained from the postulated model built with the trained sample data. Patients in the validation sample data set were excluded from the data set used to construct the training model, but the regression estimates of the latter model were used to predict the NCS outcomes for patients in the validation sample. The diagnostic performance characteristics can be used to evaluate the predictive ability of the training model if our assumption of linear and additive relationship between the test and NCS is correct. As results in panel 2 of Table 2 indicate, the validated predictive model had a sensitivity of 83%, a specificity of 73%, AUC of 78%, and an overall accuracy of 80% for the cutoff threshold of three incorrect responses. As found in the training model, there was a significant improvement in the performance

measures for the validated model when the risk factors are used in conjunction with the test to predict NCS.

3.3.2. *Superficial pain sensation*

The diagnostic performance characteristics obtained for the training model when only the superficial pain sensation modality was the predictor variable and other risk factors for neuropathy are added to the model are shown in panel 1 of Table 3. When the effect of the sensory test was considered alone, the associated values for sensitivity, specificity, AUC and overall accuracy were 48, 88, 67.9 and 55%, respectively, for a cutoff threshold of two. These values are not statistically significantly different from those obtained for the same model with the cutoff threshold of three. This model demonstrated a directly proportional association between the number of incorrect responses with the test's ability to predict the true value of NCS. The magnitude of the regression estimate for superficial pain sensation and the sensitivity and overall accuracy values for the model with the cutoff threshold of four and five incorrect responses were much larger than those obtained for the models with the cutoff threshold of two and three incorrect responses. Interestingly, when we added other risk factors for neuropathy to the model, the significant effect of superficial pain sensation disappeared. Panel 2 of Table 3 shows the results of the validated model and the associated performance characteristics when the initial training model was applied in 130 prospectively enrolled patients whose cases were not used to develop the model.

3.3.3. *Vibration by the on-off method*

Table 4 shows the diagnostic performance characteristics of both the training and validated predictive models for various cutoff thresholds for the vibration by the *on-off* method sensory test. Like the 10 g monofilament modality, the test showed a very strong positive relationship with NCS. A directly proportional association between the number of incorrect responses for vibration by the *on-off* method with the test's ability to predict the true value of NCS was observed. The effect of the screening test persisted with the addition of

Table 3

Diagnostic performance characteristics for superficial pain sensation sensory test

Model	Sensory test cutoff ^a							
	Sensory test only				Sensory test + risk factors			
	2	3	4	5	2	3	4	5
<i>Training model (n = 346)</i>								
Parameter estimate ^b	1.909 ^c	1.886 ^c	2.433 ^c	3.008 ^b	0.333	0.253	0.502	1.293
Sensitivity (%)	48	44	99	99	86	87	87	86
Specificity (%)	88	89	3	3	80	80	82	82
Accuracy (%)	55	52	80	80	85	86	86	85
Positive likelihood ratio (%)	3.95	4.11	1.02	1.02	4.39	4.41	4.77	4.73
Negative likelihood ratio (%)	0.59	0.63	0.24	0.24	0.17	0.16	0.16	0.17
AUC (%)	67.9	66.5	51.1	51.1	83.4	83.5	84.3	83.9
<i>Validated model (n = 130)</i>								
Sensitivity (%)	43	35	99	99	86	86	86	86
Specificity (%)	92	92	5	5	73	73	73	73
Accuracy (%)	57	52	71	71	82	82	82	82
Positive likelihood ratio (%)	5.30	4.38	1.04	1.04	3.18	3.18	3.18	3.18
Negative likelihood ratio (%)	0.62	0.70	0.21	0.21	0.19	0.19	0.19	0.19
AUC (%)	67.5	63.7	52.0	52.0	79.5	79.5	79.5	79.5
<i>Revised model (n = 473)</i>								
Parameter estimate ^a	2.005 ^c	1.903 ^c	2.472 ^c	2.697 ^c	0.378	0.035	0.532	0.536
Sensitivity (%)	47	42	25	23	85	86	86	86
Specificity (%)	89	90	97	98	80	80	81	81
Accuracy (%)	56	52	42	39	84	84	84	84
Positive likelihood ratio (%)	4.40	4.31	9.00	11.7	4.16	4.21	4.40	4.40
Negative likelihood ratio (%)	0.59	0.64	0.76	0.79	0.19	0.18	0.18	0.18
AUC (%)	68.2	66.1	61.7	60.4	82.2	82.7	83.1	83.1

^a Sensory test cutoff refers to the number of correct insensate responses.

^b Regression model parameter estimate of sensory test on NCS. (a) $P < 0.05$; (b) $P < 0.01$; (c) $P < 0.001$. Risk factors included in the model: neuropathy severity stage, family history of diabetes mellitus, duration of diabetes, age, height, sex, and HbA_{1c}.

risk factors in the model but was only significant for the models with cutoff thresholds of two and three. For the latter threshold, the associated sensitivity, specificity, AUC and overall accuracy were 87, 83, 85.2 and 86%, respectively.

The initial training model for this test was also applied in 130 prospectively enrolled patients whose cases were excluded from the data set used to develop the model. The resulting diagnostic performance characteristics for the validated model across the range of cutoff thresholds are shown in panel 2 of Table 4. For the model containing the sensory test and risk factors and for cutoff threshold equal to 3, we observed a sensitivity of 86%, a specificity of 83%, an AUC of 85.2% and an overall accuracy of 86%.

3.3.4. Construction of the revised interpretation model

Both the training and validation samples were combined to develop the revised model and this was used to construct the DPN risk assessment model specific to each sensory test. The latter model can be used to assess the value of each screening test as a single predictor of NCS for the diagnosis of DPN when considering the clinical and demographic characteristics of the subjects.

The diagnostic performance characteristics of the revised model for various test cutoff points are shown in the last panel of Tables 2–4 for each sensory test. In Table 2 for example, when the cutoff is three insensate responses for the SWME, the model shows a sensitivity of 86%, a specificity

Table 4

Diagnostic performance characteristics for vibration 1 (by *on-off* method) sensory test

Model	Sensory test cutoff ^a							
	Sensory test only				Sensory test + risk factors			
	2	3	4	5	2	3	4	5
<i>Training model (n = 346)</i>								
Parameter estimate ^b	2.937 ^c	3.228 ^c	3.118 ^b	2.901 ^b	1.436 ^a	1.845 ^a	2.154	1.957
Sensitivity (%)	47	44	100	100	87	87	87	87
Specificity (%)	95	97	3	3	79	83	82	82
Accuracy (%)	56	54	81	81	85	86	86	86
Positive likelihood ratio (%)	10.4	14.5	1.03	1.03	4.09	5.23	4.79	4.77
Negative likelihood ratio (%)	0.55	0.58	0.12	0.12	0.17	0.15	0.16	0.16
AUC (%)	71.3	70.4	51.3	51.3	82.8	85.2	84.5	84.3
<i>Validated model (n = 130)</i>								
Sensitivity (%)	40	34	99	99	87	86	86	87
Specificity (%)	92	97	5	5	73	73	73	73
Accuracy (%)	55	52	71	71	83	82	82	83
Positive likelihood ratio (%)	4.91	12.7	1.04	1.04	3.22	3.18	3.18	3.22
Negative likelihood ratio (%)	0.66	0.67	0.21	0.21	0.18	0.19	0.19	0.18
AUC (%)	65.8	65.9	52.0	52.0	80.0	79.5	79.5	80.0
<i>Revised model (n = 473)</i>								
Parameter estimate ^a	2.605 ^c	3.175 ^c	3.544 ^c	3.381 ^c	0.954	1.579 ^a	2.311 ^a	2.107
Sensitivity (%)	46	42	25	22	84	84	85	86
Specificity (%)	94	97	99	99	81	81	82	82
Accuracy (%)	56	54	41	39	83	83	84	85
Positive likelihood ratio (%)	7.82	14.3	26.1	23.0	4.32	4.32	4.62	4.65
Negative likelihood ratio (%)	0.58	0.60	0.75	0.78	0.20	0.20	0.18	0.17
AUC (%)	69.9	69.4	62.2	60.7	82.2	82.2	83.4	83.7

^a Sensory test cutoff refers to the number of correct insensate responses.

^b Regression model parameter estimate of sensory test on NCS. (a) $P < 0.05$; (b) $P < 0.01$; (c) $P < 0.001$. Risk factors included in the model: neuropathy severity stage, family history of diabetes mellitus, duration of diabetes, age, height, sex, and HbA_{1c}.

of 82%, accuracy of 85%, AUC of 84%, and positive and negative likelihood ratios of 4.64 and 0.18. For the superficial pain sensation test, the respective associated values are a sensitivity of 86%, a specificity of 80%, accuracy of 84%, AUC of 83%, and positive and negative likelihood ratios of 4.21 and 0.18. The observed values for vibration by the *on-off* method screening test are a sensitivity of 84%, a specificity of 81%, accuracy of 83%, AUC of 82%, and positive and negative likelihood ratios of 4.32 and 0.20, respectively.

3.3.5. Choosing the sensory test operating point

The best cut-off point on the validated model and revised model for each of the three screening tests that best accomplished the goals of this

study is less than or equal to three insensate responses (Tables 1–3). At that point, any consideration of more than three insensate responses did not produce information and power sufficient enough to discriminate between abnormal and normal neuropathy status when considering the patient's clinical and demographic characteristics.

3.3.6. The peripheral neuropathy risk assessment model

The DPN risk assessment model used the optimal operating point of greater than three insensate responses to define an 'abnormal' test result and less than or equal to three for a 'normal' test result (> 3 vs. ≤ 3) for each sensory test. The ability of each test to predict the odds of a DPN

diagnosis in patients with different clinical and demographic profiles was evaluated. Furthermore, the specific sensory test modality that best predicts the odds of DPN, as well as the clinical characteristic risk factors were identified. The degree to which the power of each screening test to predict a diagnosis of DPN changed when adjusted for the effects of risk factors in the model was established. Table 5 illustrates the results of the logistic regression model used to address these issues.

For the SWME sensory test, the odds of DPN are positively and significantly associated with the sensory test, age, HbA_{1c}, and duration of diabetes, and significantly linearly related with neuropathy severity stage. The odds ratio of 3.149 ($= e^{1.47}$, $P < 0.01$) is the odds of an abnormal NCS finding for an abnormal SWME test result relative to normal test results adjusted for neuropathy severity stage and other risk factors. The estimated odds compared with the adjusted odds of 7.973 ($= e^{2.076}$, $P < 0.001$, see the parameter estimate for the revised model in Table 2), suggest that the

effect of the SWME test in detecting DPN was more than halved when demographic and clinical factors were added into the model.

The estimated values for neuropathy severity stage indicate a linear relationship with DPN, abnormal NCS findings are more prevalent in the severe neuropathy severity stage group, followed by the moderate, mild, and non-neuropathic groups, respectively. The least probable likelihood of abnormal NCS results is found in the reference subjects. As seen in the Dyck et al. study, [21] abnormality as indicated by NCS differentiates non-diabetic control subjects from subjects with diabetes, as well as subjects with diabetes with and without neuropathy. The values of 1.063 ($= e^{0.061}$) and 1.076 ($= e^{0.073}$) indicate the relative change in the odds of having DPN for each unit year increase in age and unit year increase in the duration of diabetes, respectively, holding other variables constant. The odds of DPN also increase by a factor of 1.284 per percent increase in HbA_{1c}.

Table 5

Regression estimates from final revised model for predicting neuropathy using the three sensory tests^a

Variable	Screening test model					
	10 g Monofilament		Superficial pain sensation		Vibration 1 (by on-off method)	
	Estimate	Odds ratio	Estimate	Odds ratio	Estimate	Odds ratio
Intercept	-3.528	0.029	-3.940	0.019	-2.432	0.088
Sensory test (abnormal versus normal) ^a	1.147	3.149 ^b	0.035	1.036	1.579	4.850 ^a
<i>Neuropathy severity stage</i>						
Reference	-3.199	0.041 ^c	-3.715	0.024 ^c	-3.506	0.030 ^c
Non-neuropathic	-2.697	0.067 ^c	-3.169	0.042 ^c	-2.938	0.053 ^c
Mild	-1.753	0.173 ^a	-2.039	0.130 ^b	-1.840	0.159 ^b
Moderate	-1.076	0.341	-1.269	0.281	-1.228	0.293
Severe (reference)						
Age (year)	0.061	1.063 ^c	0.067	1.069 ^a	0.060	1.062 ^c
Height (cm)	0.419	1.520	1.092	2.980	0.259	1.296
Sex (female vs. male)	-1.152	0.316	-0.994	0.370 ^a	-0.996	0.037 ^a
Duration of diabetes (year)	0.073	1.076 ^c	0.070	1.073	0.067	1.068 ^b
Family history of diabetes (none vs. yes)	0.636	1.889	0.552	1.737	0.509	1.664
HbA _{1c}	0.250	1.284 ^a	0.209	1.232	0.204	1.226

^a Regression model parameter estimate of sensory test on NCS. A sensory test cutoff of three insensate responses was used to differentiate between 'abnormal' from 'normal' screening test result (see text for detailed explanation). (a) $P < 0.05$; (b) $P < 0.01$; (c) $P < 0.001$.

Similar results were observed for vibration by the *on-off* method screening test. The odds of abnormal NCS findings for patients with abnormal test results relative to normal test results adjusted for risk factors is 4.850. Neuropathy severity stage, age, duration of diabetes, and sex were all significantly associated with the NCS outcome. The odds ratio for the gender variable, 0.369 ($= e^{-0.996}$), indicates that abnormal NCS outcomes are less likely for females compared with males in this population. Similar to the SWME test, inclusion of other risk factors in the model leads to more than a 50% reduction in the predictive ability of this test to diagnose DPN. When the risk factors are excluded from the model, the odds of a DPN diagnosis are 23.93 ($= e^{3.1735}$, $P < 0.001$) times more likely to occur among patients with abnormal test results relative to those with normal test results (see the last panel of Table 4).

The findings for superficial pain sensory test differ from those of the two modalities discussed above. Although NCS results were positively associated with this test when keeping all other factors in the model constant, the effect was not significant. On the other hand, the observed ability of superficial pain sensation (odds = 6.593, $P < 0.001$) to predict NCS outcomes was significant when the association between this test and NCS alone were analyzed (see Table 3).

4. Discussion

Nerve conduction studies are valuable as surrogate markers or endpoint variables for peripheral neuropathy in controlled clinical trials [2,5,22]. The measures are very accurate, objective, sensitive, and also a valid instrument for the study of severity and progression of neuropathy [1,2]. While NCS measures have been used successfully in large, multi-center clinical investigations of DPN, the use of NCS for general screening for neuropathy in the diabetes clinics has been a subject of debate [3,4,8]. The common criticisms against the use of this measure in the primary care clinic or diabetes clinic

are related to the time-consuming and costly nature of NCS, as well as to limited availability.

Given these limitations of NCS, other screening methods for neuropathy in the diabetes clinic are desirable. The ideal neuropathy screening instrument is one that provides comparably accurate and objective results, but that is less time-consuming, readily available, and reliable. In terms of statistical properties, the screening instrument must have high sensitivity, specificity and positive likelihood ratio with a low negative likelihood ratio to be a valid instrument. When such an instrument is applied to a broad spectrum of subjects of unknown neuropathy and diabetes status, the method must possess sufficient statistical power to clearly identify non-diabetic control subjects, as well as to determine the neuropathic status of subjects with diabetes.

Rigorous investigation has demonstrated the utility of the SWME as a screening instrument for neuropathic foot complications [23,24]. The current study applied similar rigorous investigation to simple clinical tests in comparison to NCS, and reports that both the SWME and vibration by the *on-off* method meet the statistical requirements for a valid screening instrument as assessed by a cross-sectional study design. Sensation to superficial pain has less power to discriminate between neuropathic and non-neuropathic subjects, although this test met some of the statistical requirements for screening. Within the context of operating point thresholds considered in this study, the SWME and vibration by the *on-off* method tests demonstrate acceptable diagnostic performance characteristics in terms of high sensitivity and specificity, as well as high positive and low negative likelihood ratios. The diagnosis of DPN, as evaluated by this study, defines less than or equal to three insensate responses as the operating point threshold for both sensory tests. More than three incorrect responses to the stimuli rule in peripheral neuropathy. This testing is brief, logistically simple, inexpensive, and readily available in contrast to NCS [17]. However, this specific operating point requires the knowledge of DPN clinical risk factors, and the ability to incorporate these vari-

ables into the interpretation of physical examination results.

The earlier published initial report on the validity of three sensory screening modalities in the same study cohort [17] revealed that the SWME, vibration by the *on-off* method, as well as superficial pain sensation are rapid, reliable, reproducible, and valid tests with similar statistical operating characteristics on logistic regression analyses. The discovery of five or more insensate stimuli out of eight was found in that investigation to rule in DPN. The current study was intended to extend the statistical inferences of the initial report by applying a more sophisticated statistical modeling technique in order to further demonstrate the validity of these maneuvers. The model was derived from a training set, and tested in a validation set with an emphasis on incorporating potential DPN clinical risk factors. This study demonstrates that the knowledge of clinical risk factors alters the operating point for sensory tests in diagnosing DPN. This finding, since it is derived from a separate statistical methods approach to the interpretation of a common cohort and provides similar results, increases the confidence in these sensory tests as valid screening maneuvers for DPN. Since a screening maneuver is performed, by definition, in the absence of the knowledge of pre-test probability for the diagnosis of DPN, it is recommended that the operating points derived from an unadjusted analysis be used in clinical practice. Therefore, we recommend annual screening with either the SWME or vibration by the *on-off* method as earlier described [17] by the technique of five or more insensate responses out of eight for the conditional diagnosis of DPN in the primary care and diabetes clinic.

References

- [1] Proceedings of a consensus development conference on standardized measures in diabetic neuropathy. *Neurology* 42 (1992) 1823–1839.
- [2] Diabetes Control and Complications Trial (DCCT) Research Group, Effect of intensive diabetes treatment on nerve conduction in the diabetes control and complications trial, *Ann. Neurol.* 38 (1995) 869–880.
- [3] V. Bril, Current treatment of diabetic neuropathy, *Can. J. Diab. Care* 24 (2000) 67–75.
- [4] P. Dyck, K. Kratz, K. Lehman, et al., The Rochester Neuropathy Study: design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests, *Neurology* 41 (1991) 799–807.
- [5] D.R. Cornblath, V. Chaudhry, K. Carter, et al., Total neuropathy score: validation and reliability study, *Neurology* 53 (8) (1999) 1660–1664.
- [6] E.L. Feldman, M.J. Stevens, Clinical testing in diabetic peripheral neuropathy, *Can. J. Neurol. Sci.* 21 (4) (1994) S3–S7.
- [7] American Diabetes Association, Report and recommendations of the San Antonio conference on diabetic neuropathy, *Diabetes* 37 (1988) 1000–1004.
- [8] D. Armstrong, L.A. Lavery, S. Vela, T. Quebedeaux, J. Fleischli, Choosing a practical screening instrument to identify patients at risk for diabetic foot ulceration, *Arch. Intern. Med.* 158 (70) (1998) 63–66.
- [9] M.S. Rendell, J.J. Katims, R. Richter, F. Roland, A comparison of nerve conduction velocities and current perception thresholds as correlates of clinical severity of diabetic sensory neuropathy, *J. Neurol. Psychiatry* 52 (4) (1989) 502–511.
- [10] P.J. Dyck, P.C. O'Brien, Meaningful degrees of prevention or improvement of nerve conduction in controlled clinical trials of diabetic neuropathy, *Diab. Care* 12 (1989) 649–652.
- [11] Diabetes Control and Complications Trial (DCCT) Research Group, The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus, *New Engl. J. Med.* 329 (1993) 977–986.
- [12] A. Boulton, Why bother educating the multidisciplinary team and the patient — the example of prevention of lower extremity amputation in diabetes, *Patient Educ. Couns.* 26 (1995) 183–188.
- [13] A. Boulton, Lowering the risk of neuropathy, foot ulcers and amputations, *Diab. Med.* 15 (suppl 4) (1998) S57–S59.
- [14] S. Rith-Najarian, T. Stolusky, D. Gohdes, Identifying diabetic patients at high risk for lower extremity amputation in a primary health care setting, *Diab. Care* 15 (1992) 1386–1389.
- [15] B.A. Perkins, The clinical assessment of neuropathy in the diabetes clinic: background information for the design of a methodologically sound clinical trial, *Can. J. Diab. Care* 24 (Supp 2) (2000) 61–66.
- [16] D.M. Huse, G. Oster, A.R. Killen, M.J. Lacey, G.A. Colditz, The economic costs of non-insulin dependent diabetes mellitus, *J. Am. Med. Assoc.* 262 (1989) 2708–2713.
- [17] B.A. Perkins, D. Olaleye, B. Zinman, V. Bril, Simple screening tests for peripheral neuropathy in the diabetes clinic, *Diab. Care* 24 (2001) 250–256.
- [18] SAS Institute Inc. SAS/STAT software: changes and enhancements, version 6 through version 6.11. Cary, NC: SAS Institute, 1996.

- [19] Diabetes Control and Complications Trial (DCCT) Research Group, Factors in development of diabetic neuropathy: baseline analysis of neuropathy in feasibility phase of DCCT, *Diabetes* 37 (1988) 476–481.
- [20] J.A. Hanley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology* 143 (1982) 29–36.
- [21] P.J. Dyck, J.L. Karnes, J. Daube, P.C. O'Brien, F.J. Service, Clinical and neuropathological criteria for the diagnosis and staging of diabetic polyneuropathy, *Brain* 108 (Pt. 4) (1985) 861–880.
- [22] P.J. Dyck, L.J. Melton III, P.C. O'Brien, F.J. Service, Approaches to improve epidemiological studies of diabetic neuropathy: insights from the Rochester Diabetic Neuropathy Study, *Diabetes* 46 (Suppl. 2) (1997) S5–S8.
- [23] M. Smieja, D. Hunt, D. Edelman, E. Etchells, J. Cornuz, D. Simel, Clinical examination for the detection of protective sensation in the feet of diabetic patients: international cooperative group for clinical examination research, *J. Gen. Intern. Med.* 14 (1999) 418–424.
- [24] D. Armstrong, L.A. Lavery, S. Vela, T. Quebedeaux, J. Fleischli, Choosing a practical screening instrument to identify patients at risk for diabetic foot ulceration, *Arch. Intern. Med.* 158 (1998) 289–292.