AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 39:361–368 (2001)

Agreement Between Clinical Examination and Quantitative Tests of Neurologic Function Among 384 Subjects

Kyle Steenland,* Robert Dick, and Lawrence Fine

Background *Quantitative neurological tests are often cheaper and easier than clinical examinations, and provide continuous data which may discriminate between exposed and nonexposed groups with more sensitivity than dichotomous (normal/abnormal) examination data.*

Methods *We compare clinical examinations and analogous quantitative tests for arm tremor, postural sway, and vibrotactile sensitivity (finger and toe), for 384 subjects.*

Results *The "abnormal" clinical outcomes studied were relatively common (range, 3–36%), and did not result in impairment of daily activity for affected subjects. All the quantitative tests were reasonably good predictors of the corresponding clinical outcome. The most predictive test was for toe vibrotactile sensitivity. The probability of an abnormal clinical result for those in the worst quartile for the toe test was 0.63, compared with 0.36 for all subjects.*

Conclusions *Our results suggest that certain quantitative tests might be used in epidemiologic studies instead of a physical examination.* Am. J. Ind. Med. 39:361–368, 2001. *Published 2001 Wiley-Liss, Inc.*†

KEY WORDS: neurologic tests; epidemiology; neurologic exam

INTRODUCTION

We recently completed an epidemiologic study of 193 termiticide applicators and 191 nonexposed controls [Steenland et al., 2000]. The focus of the study was on neurologic function. Subjects were examined by a clinical neurologic examination and a variety of quantitative tests. The clinical examination and the quantitative tests were analogous for three endpoints: (1) sensitivity to vibration in the toe and finger (peripheral nerve function); (2) tremor (central nerve function), and (3) sway (both peripheral and central nerve function). Here, we report the relationship between findings on the clinical examination, which were usually scored as "normal/abnormal," and the quantitative tests, which were generally scored as continuous variables.

The question of interest is the degree to which findings on the quantitative tests can predict findings on the clinical examination. If these tests are good predictors of a clinical abnormality, they might be used in place of a clinical examination in epidemiologic field studies. It should be noted that a finding a clinical abnormality for the outcomes here was not uncommon and did not necessarily imply daily impairment nor was sufficient, by itself, to make a diagnosis.

The quantitative tests discussed here offer several advantages over a clinical examination. One advantage is that the quantitative tests provide a result which is a continuous rather than a dichotomous variable, and, therefore, allow for detecting more subtle neurologic differences between groups (e.g., exposed vs. nonexposed subjects in an epidemiologic study) than does the clinical examination, assuming the test and clinical examination are measuring the same phenomenon. They also may be considerably less expensive

National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226-1998

*Correspondence to: Kyle Steenland, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226-1998. E-mail: nsteenland@cdc.gov

Accepted 8 December 2000

and faster to perform than a clinical examination, especially in field studies with large numbers of subjects. They are easily automated and less subject to error or bias on the part of the tester, relative to the clinician.

On the other hand, the quantitative tests do not yet have validated norms among nonexposed populations, although this situation is improving. A clinical test, in contrast, does by definition have such norm. It usually divides subjects into ''normal'' and ''abnormal,'' albeit based on a qualitative judgment of the neurologist rather than a quantitative criterion.

If the quantitative tests were good predictors of the clinical examination, they might be used in place of clinical examinations to differentiate between exposed and nonexposed groups in epidemiologic studies, and might also be used as a screening tool to refer some subjects to subsequent clinical testing. There exist only a few prior studies of this subject (see Discussion).

METHODS

Details of the original cross-sectional study of 193 fermiticide applicators and 191 controls can be found in Steenland et al. [2000]. Briefly, the focus of the study was neurologic function, possibly related to exposure to pesticide. Applicators were chosen from a state list of applicators in central North Carolina. Nonexposed controls were approximately evenly divided between either blue-collar workers (North Carolina state employees) or friends of the exposed. Tests were conducted on weekends at a central location. All participants lived within 2 hr of the testing site. Participants spent 4–5 hr going through a battery of neurologic examinations. The clinical examination evaluated selected cranial nerves, speech, resting and postural tremor, coordination and strength, sensation, reflexes, station, and gait. Quantitative tests included sensory, motor, and cognitive tests. All subjects received US $250 compensation for their time. The study was approved by a human subject review board; all subjects signed consent forms. All subjects were male except six females, who did not differ in outcomes from males and were retained in the analysis. No significant exposure-related effects were found for the tests considered here, either clinical examinations or quantitative tests. We, therefore, considered the combined exposed and nonexposed population for the present analysis. In the original study, we deleted two applicators without good work history and two controls who had worked as applicators. However, these four subjects were retained in the present analysis.

The Clinical Examination

A clinical neurologic examination was performed by two neurologists. Study subjects were randomly examined by either one or the other neurologist, but not by both. Each neurologist examined about half of the study subjects. Prior to beginning the study, both neurologists performed several practice examinations on study investigators and discussed their findings to try to make their approach uniform. To account for any differences between neurologist, we included an indicator variable for neurologist in our regression analyses, which was tantamount to doing an analysis stratified by neurologist.

The neurologists did not know the results of the corresponding quantitative test. Similarly, the staff who conducted the quantitative tests did not know the results of the corresponding clinical examination. Furthermore, the workers also did not know their own results for any quantitative test or clinical examination at the time these procedures were performed. In addition, neither neurologist or testers knew the exposure status of participants.

The neurologists performed three common examinations [Degowin and Degowin, 1981] which had analogous quantitative tests.

(a) *Vibrotactile sensitivity.* Examination for sensitivity to vibration for big toe and index finger (right and left), using a tuning fork vibrating at 125 Hz. Each of the four outcomes was coded as normal or abnormal, based on whether the subject reported feeling the vibrating fork. Findings between right and left digits were almost always concordant (96% for toe, 98% for finger). In our analysis we combined findings for right and left digits; a finding of abnormal on either digit was counted as abnormal.

(b) *Tremor.* Examination for upper extremity tremor, using right and left arms. Results were scored as none, mild, moderate, and severe, depending on the clinical judgment of the neurologist. No subjects had severe tremor. Again, results for right or left arm were virtually always concordant (98%), and we combined results for left and right arms; the highest score for either arm was used as the score in the analysis. Analyses considered either a three-category ordinal outcome, or a dichotomous outcome, none versus mild/moderate.

(c) *Sway.* The Romberg examination for sway consists of comparing sway while standing with eyes open versus eyes closed. This exam was scored as normal or abnormal, depending on the clinical judgment of the neurologist.

Quantitative Tests

Below we list the quantitative tests which corresponded to the clinical examinations.

(a1) *Vibrotactile sensitivity.* Test of sensitivity for vibration for the big toe and index finger (nondominant). The fully automated test (Vibrometry System-Brüel & Kjær) was conducted at two vibration frequencies (31.5 Hz and 125 Hz) over an intensity range from 90 to 160 dB using the Békésy method [Békésy, 1947]. Decibels are defined as

$20 \log_{10} (a_{rms}/a_{ref})$, where a_{rms} is the root mean square of the vibration acceleration, and a_{ref} a reference vibration acceleration of 10^{-6} m/s^2. In the analysis, we considered both frequencies separately. The test produces a mean threshold for detection at each frequency in decibels, which was used for the analysis. This test corresponded directly to the clinical examination for toe and finger sensitivity to vibration.

(a2) *Nerve Conduction velocity of the peroneal motor and sural sensory nerves*. This test was compared to the clinical examination for the toe, since both are measures of peripheral nerve function in the lower limbs. Nerve conduction velocity is measured in meters per second. For sensory nerves the meters measure the distance between stimulating and recording electrodes, and the time measures time elapsed between stimulation of the nerve and recording of the response (sensory nerve action potential). For motor nerves, two stimuli are given at two different points and two times measured (time from stimulation to compound motor action potential). Peak amplitude (the size of the maximal response of the action potential) was also recorded. All tests were done using surface electrodes, at 33.0°C. Sensory impulses were antidromic (reverse of normal direction).

(b) *Tests of arm tremor.* This test corresponded directly to the test of upper extremity tremor on the clinical examination. Tremor was measured by a NIOSH-developed handheld device which measured tremor via accelerometers [Galinsky et al., 1990]. The instrumentation has the capability to detect both visible (approximately 1–6 Hz) and nonvisible (approximately 7–30 Hz) tremor. The test is microcomputer controlled and produces a power-spectrum (a measure of the amplitude of tremor) for each tremor frequency measured. Analyses considered horizontal tremor and vertical tremor of the dominant arm, averaged across all frequencies for each minute of a 3-min test. Here we report analyses for total horizontal tremor measured during the second minute [the log (base 10) of the sum across all frequencies], which proved to be the best predictor of tremor on the clinical examination.

(c) *Tests of sway.* This test corresponded directly to the Romberg test of sway on the clinical examination. Postural stability was measured during 30-sec tests with a microcomputer controlled force platform (Advanced Mechanical Technology Instruments) using varied test conditions [Battacharya et al., 1995; Dick et al., 1999]. We used the ratio of results for eyes closed to eyes open on a hard platform, which was the best predictor of the analogous Romberg clinical exam. We used the outcome of the ratio of length of sway with eyes closed to eyes open on a hard platform. This outcome was the most predictive of the clinical examination result.

(3) *Analytic approach.* Analyses included of univariate *t*-tests comparing the means of the quantitative tests by category of the clinical examination (e.g., normal vs. abnormal). Subsequent analyses consisted of logistic regressions in which the outcome was the clinical examination (normal/abnormal), and the predictor variables were the results of the quantitative tests, age, and neurologist conducting the examination. The latter two variables were significantly related to most outcomes and were included as potential confounders. The variable for the quantitative tests was analyzed either as a continuous or categorical (quartile) variable. The test of significance of the continuous variable was taken as a test for trend (linear trend in the log odds). Other potentially confounding variables (e.g., exposure status, smoking, body mass index, race) were not predictors of clinical examination results and were not included in the model. Interaction terms between quantitative test and either age or neurologist were tested but were not statistically significant, with the exception of age and peroneal nerve conduction in predicting toe vibrotactile sensitivity. For this clinical outcome a stronger predictive effect of peroneal nerve conduction was seen in older subjects; we report results separately for those over age 40.

Finally, we have also calculated the sensitivity, specificity, and (1-specificity) at four cutpoints (25th, 50th, 66th, and 75th percentiles) for four of the most predictive quantitative tests, under the assumption that the clinical test is the ''gold standard.'' A positive quantitative test is here defined as one above the cutpoint. Sensitivity is defined as the percentage of positive tests among all those positive (abnormal) on the clinical exam. Specificity is defined as the percentage of negative tests among all those negative (normal) on the clinical examination. Sensitivity and (1-specificity) at different cutpoints are sometimes used to construct ''receiver–operator characteristic'' or ''ROC'' curves, which are graphs of sensitivity versus (1-specificity) with cutpoints dividing ''abnormal'' versus ''normal'' tests ranging across all the values of the predictor values. We have also calculated the positive predictive power (percent abnormal on examination out of all abnormal on quantitative test) for the 75th percentile cutpoint.

RESULTS

Table I shows the results of *t*-tests comparing the means of the quantitative tests for those normal or abnormal on the clinical examination. Inspection of Table I shows that these means were significantly different from each other at the $P < 0.05$ level, with the exception of those for the Romberg examination which was limited by a small sample size (only 12 subjects were found to be abnormal by the clinicians), and for vertical tremor. These significant differences would be expected assuming the quantitative tests had some predictive power for the outcomes on the examination. These *t*-tests do not take into account potential confounding by age or neurologist performing the examination.

TABLE I. Univariate Results of Clinical and Quantitative Tests

Quantitative test	Clinical examination	No. of normals on exam (%)	Normals: test mean (SD)	No. of abnormals on exam (%)	Abnormals: test mean (SD)	*P*-value, *t*-test difference in test means
Toe vibration (31.5 Hz) (dB)	Toe vibration	245 (64)	111.5 (8.0)	139 (36)	116.2 (9.9)	0.0001
Toe vibration (125 Hz) (dB)	Toe vibration	245 (64)	113.9 (11.2)	139 (36)	123.0 (13.8)	0.0001
Peroneal velocity (m/sec)	Toe vibration	245 (64)	47.2 (4.1)	139 (36)	45.4 (4.0)	0.0001
Peroneal amplitude, fibers ennervated	Toe vibration	245 (64)	7.93 (2.57)	139 (36)	7.16 (2.95)	0.01
Sural velocity (m/sec)	Toe vibration	245 (64)	47.5 (4.2)	139 (36)	45.9 (4.1)	0.0004
Sural amplitude, fibers ennervated	Toe vibration	245 (64)	20.0 (6.7)	55 (14)	18.1 (6.8)	0.007
Finger vibration (31.5 Hz) (dB)	Finger vibration	329 (86)	105.9 (8.0)	55 (14)	109.6 (10.3)	0.01
Finger vibration (125 Hz) (dB)	Finger vibration	329 (86)	104.1 (9.9)	55 (14)	107.5 (11.5)	0.02
Horizontal arm tremor, units of power spectrum[a]	Arm tremor	289 (75)	4.77(0.53)	95 (25)	4.95 (0.52)	0.005
Vertical arm tremor, units of power spectrum	Arm tremor	289 (75)	4.73 (0.57)	95 (25)	4.80 (0.39)	0.14
Ratio of length of sway, eyes closed to eyes open	Romberg exam for sway	372 (97)	1.49 (0.30)	12 (3)	1.69 (0.41)	0.03

[a]Arm tremor abnormal on clinical examination groups moderate (n = 13)and mild (n = 82) arm tremor. The test means for these two groups for horizontal tremor were 5.38 (0.69) and 4.88 (0.46), respectively, while the means for vertical tremor were 5.25 (0.46) and 4.74, respectively.

Tables II–V shows the results of the multivariate logistic regression analyses in which the results of the quantitative tests are used to predict the outcome of the clinical examination, while controlling for age and neurologist performing the examination.

Table II shows that peroneal nerve conduction velocity and amplitude were statistically significant predictors (based on the *P*-value for trend) of an abnormal clinical examination of toe vibrotactile sensitivity. The relationship between test and examination was strongest for older subjects.

Quantitative vibration sensitivity tests for the toe are directly analogous to the clinical test, and show a strong relationship with it in Table II, at either 31.5 Hz or 125 Hz (these two quantitative tests are highly correlated, $r = 0.79$).

Table III shows the results for an abnormal examination for finger vibrotactile sensitivity, compared with the quantitative test for finger vibrotactile sensitivity at 31.5 and 125 Hz. These two tests are correlated with each other (Pearson's correlation coefficient, 0.61). There are significant trends of worse performance on the test predicting the

TABLE II. Logistic Regression*, Odds of an Abnormal Examination for Toe Vibrotactile Sensitivity by Nerve Conduction Velocity and Amplitude, and by Vibrotactile Sensitivity (Quantitative Test)

Quantitative test	Odds ratio of abnormal examination by quartile (*P*-value)	*P*-value for trend[a]
Peroneal nerve velocity	1.00, 0.72 (.39), 1.46 (.33), 1.66 (.19)	0.009
Peroneal amplitude	1.00, 1.58 (.24), 1.54(.26), 2.64 (.01)	0.02
Peroneal nerve velocity, age 40+	1.00, 0.45 (.20), 1.56 (.45), 2.22 (.15)	0.003
Peroneal amplitude, age 40+	1.00, 1.88 (.29), 1.65(.36), 3.95 (.01)	0.008
Sural nerve velocity	1.00, 0.97 (.92), 1.80 (.12), 1.63 (.24)	0.09
Sural amplitude	1.00, 0.69(.33), 1.21 (.61), 1.00 (.99)	0.98
Toe vibrotactile sensitivity, 31.5 Hz	1,00, 1.63 (.24), 1.31 (.50), 3.82 (.002)	0.002
Toe vibrotactile sensitivity, 125 Hz	1.00, 1,56 (.27), 1.80 (.15), 5.03 (.0001)	0.0001

*Controlled for age, neurologist performing test.
[a] Test of coefficient for continuous variable for toe vibrotactile sensitivity.

TABLE III. Logistic Regression*, Odds of an Abnormal Examination for Finger Vibrotactile Sensitivity on Quantitative Examination

Quantitative test	Odds ratio of abnormal examination by quartile (*P*-value)	*P*-Value for trend[a]
Finger vibrotactile sensitivity, 31.5 Hz	1.00, 1.85 (.24), 3.85 (.006), 3.01 (.02)	0.004
Finger vibrotactile sensitivity, 125 Hz	1.00, 2.69 (.06), 2.30 (.10), 3.07 (.02)	0.03

*Controlled for age, neurologist performing test.
[a]Test for coefficient for continuous variable for finger vibrotactile sensitivity.

TABLE IV. Logistic Regression*, Odds of an Abnormal Examination for Upper Extremity Tremor by Arm Tremor on Quantitative Test

Quantitative test	Odds ratio of abnormal examination by quartile (*P*-value)	*P*-Value for trend[a]
Horizontal tremor during 2nd minute	1.00, 2.50 (.02), 2.50 (.02), 4.68 (.02)	0.002
Vertical tremor during 2nd minute	1.00, 1.22 (.60), 2.27 (.02), 2.29 (.02)	0.11

*Controlled for age, neurologist performing test.
[a] Test for coefficient for continuous variable for tremor.

clinical outcome, but quartile analysis does not indicate monotonically increasing odds ratios. It would appear that subjects in all three of the upper quartiles of the finger vibrotactile quantitative test have a two to threefold risk of an abnormal clinical examination, suggesting lack of a strong correspondence between the test and the examination.

Table IV shows the results for an abnormal examination for arm tremor (mild and moderate combined) as predicted by the quantitative tests for horizontal and vertical tremor in the 2nd minute of a 3-min test. Horizontal tremor showed a significant positive trends.

Table V shows the results for an abnormal Romberg test on clinical examination, versus the best predictor of the quantitative test for postural sway (ratio of length of sway, eyes closed vs. eyes open, hard surface). Odds ratios by quartile of the quantitative test do not indicate an monotonically increasing trend. The predictive value of the quantitative test is inherently difficult to estimate because of the low number (n = 12) of abnormals on the clinical Romberg test.

Table VI shows the sensitivity, specificity, (1-specificity), and positive predictive value for four of the tests which were among the better predictors of clinical outcomes,

TABLE V. Logistic Regression*, Odds of an Abnormal Romberg (Sway) Examination by Postural Sway Results (Quantitative Test)

Quantitative test	Odds ratio of abnormal examination by quartile (*P*-value)	*P*-Value for trend[a]
Ratio of length of sway, eyes closed to eyes open, hard surface	1.00, 0.99 (.99), 3.78 (.24), 2.46 (.43)	0.08

*Controlled for age, neurologist performing test.
[a] Test for coefficient for continuous variable for ratio of sway.

TABLE VI. Sensitivity, Specificity, and (1-Specificity) at 25th, 50th, 66th, and 75th Percentiles for Four Quantitative Tests Most Predictive of an Abnormal Result on Clinical Examination*

Test, percentile cutpoint[a]	Clinical examination	Sensitivity	Specificity	1-Specificity	Positive predictive value
Toe vibration, 75%	Toe vibration	0.45	0.85	0.15	0.63
Toe vibration, 66%	Toe vibration	0.50	0.70	0.30	
Toe vibration, 50%	Toe vibration	0.68	0.55	0.45	
Toe vibration, 25%	Toe vibration	0.87	0.29	0.71	
Peroneal velocity, 25%[b]	Toe vibration	0.35	0.80	0.20	0.50
Peroneal velocity, 33%	Toe vibration	0.46	0.74	0.26	
Peroneal velocity, 50%	Toe vibration	0.63	0.58	0.42	
Peroneal velocity, 75%	Toe vibration	0.82	0.29	0.71	
Finger vibration, 75%	Finger vibration	0.35	0.76	0.24	0.19
Finger vibration, 66%	Finger vibration	0.47	0.71	0.29	
Finger vibration, 50%	Finger vibration	0.65	0.51	0.49	
Finger vibration, 25%	Finger vibration	0.82	0.25	0.75	
Arm tremor, 75%	Arm tremor	0.37	0.79	0.21	0.36
Arm tremor, 66%	Arm tremor	0.43	0.70	0.30	
Arm tremor, 50%	Arm tremor	0.64	0.55	0.45	
Arm tremor, 25%	Arm tremor	0.87	0.29	0.71	

*Sensitivity, percent abnormals by test among all abnormals on exam; specificity, percent normals by test among all normals on examination; positive predictive value, percent abnormals on exam among all abnormal on test (calculated only for most predictive quartile).
[a]Finger test using 31.5 Hz, toe test using 125 Hz, temor test using horizontal tremor. Values above cutpoint define an "abnormal" test, except for peroneal amplitude where values below the cutpoint define an "abnormal" test.
[b]Worse peroneal velocity was slower velocity, so the ordering of categories is reversed.

evaluated at four different cutpoints (25th, 50th, 66th, and 75th percentiles). For each cutpoint the quantitative test was defined as ''positive'' if the test result exceeded the cutpoint (with the exception of peroneal nerve amplitude, where the reverse was true). This table assumes that the clinical test is the ''gold standard.'' Table VI shows the usual trade-off of sensitivity and specificity. The degree to which sensitivity generally exceeds (1-specificity) is one measure of the predictive ability of the test for the clinical outcome, the degree to which the test correctly predicts true abnormals versus the degree to which it incorrectly predicts abnormals from normals. For example, consider toe vibrotactile sensitivity test as the predictor of a clinical finding of abnormal toe sensitivity. Using the median value of the test as the cutoff, this test correctly identifies as abnormal two-thirds of all subjects; nonetheless it also has a 45% false positive rate at this cutoff point. Another measure of predictive ability is the positive predictive power, the percent of those classified as abnormal on the quantitative test who were also abnormal on the clinical exam. This measure is notably high only for the toe vibrotactile sensitivity.

Table VI illustrates the predictive power of the quantitative test, without reference to age. The logistic regressions with both age and quantitative test into the model can potentially provide additional useful information, especially when age is an important predictor of the clinical exam result in addition to the quantitative test. Age is a highly significant predictor of all examinations in Table VI except sensitivity to finger vibration. Table VII gives the logistic model results (with age and test in the model) for the same tests/examination as Table VI. Also given are the sensitivity, specificity, and predictive power of a positive test, based on using the 75th percentile of the predicted probabilities from the logistic model. The models in Table VII do not include a dichotomous variable for neurologist as in the earlier tables, as that variable was specific to our own study and not generalizable. Table VII does not indicate, in general, much change from Table VI, although the addition of age to the model provides some improvement for tests for an abnormal toe exam.

DISCUSSION

We have evaluated the predictive ability of quantitative tests of neurologic function for results of a clinical neurologic examination. The single best predictor was the vibrotactile sensitivity test for predicting abnormal toe vibrotactile sensitivity on clinical examination. Peroneal nerve conduction was also a good predictor of toe abnormality on clinical exam.

TABLE VII. Logistic Model Results for Selective Quantitative Tests and Examinations

Test	Clinical examination	Logistic model for log odds of abnormal exam	Sensitivity at 75th percentile[a]	Specificity at 75th percentile[a]	Positive predictive value at 75th percentile[a]
Toe vibration	Toe vibration	−8.85+0.068 age+ 0.046 toe test	0.48	0.88	0.72
Peroneal velocity	Toe vibration	−0.228+.082 age− 0.080 velocity	0.48	0.88	0.68
Finger vibration[b]	Finger vibration	−5.53+0.013 age+ 0.028 finger test	0.31	0.75	0.17
Arm tremor[b]	Arm tremor	−5.50+0.039 age+ 0.582 tremor	0.36	0.79	0.36

*Sensitivity, percent abnormals by test among all abnormals on examination; specificity, percent normals by test among all normals on examination; positive predictive value, percent abnormals on examination among all abnormal on test (calculated only for most predictive quartile). The cutpoint for determining sensitivity, specificity, and positive predictive value was the 75th percentile of the predicted probabilities from the logisitic model.
[a]Finger test using 31.5 Hz, toe test using 125 Hz, temor test using horizontal tremor

The finding of a stronger relationship between the peroneal nerve conduction test and an abnormal clinical finding for toe vibrotactile sensitivity for older people versus younger ones may be related to the fact that nerve conduction velocity is fairly constant until age 50, when it begins to decrease [Davis-King et al., 1992]. Clinical dysfunction may occur only after the number of healthy neurons drops below some critical number (a threshold effect).

The weaker correspondence between quantitative test and clinical examination for index finger tactile sensitivity, compared with great toe vibrotactile sensitivity, may be related to the fact that the sensory neurons of the lower limb are more vulnerable to damage because of their greater length, compared to the shorter length of the neurons in the upper limb. In our study population the prevalence of an abnormal clinical outcome for the toe was 36%, compared with 14% for the finger. The standard deviation of the quantitative test for toe vibrotactile sensitivity (at 125 Hz) was 13.0, compared 10.2 for the comparable statistic for the finger. The greater variation in the quantatitive test for the toe compared to the finger, and the greater prevalence of an abnormal clinical outcome for the toe compared to the finger, may underlie the more pronounced correspondence between test and clinical outcome for the toe versus the finger.

The weak relationship between the quantitative test for postural sway and the clinical Romberg examination may be partly a product of limited sample size; there were only 12 abnormal Romberg examinations.

It should be noted that we have here considered the clinical examination as a ''gold standard'' whereas in fact the decision by the two clinicians in this study to label a subject as normal or abnormal was a subjective judgment which necessarily involved some error in relation to some unknown true classification based on current clinical practice across all neurologists. The degree to which the clinicians in this study were themselves inaccurate will tend to reduce the true predictive ability of the quantitative tests.

Our results are consistent with findings from the few other studies on this issue. Dick et al. [1997] found good agreement between an arm tremor score in the upper 20% on the same quantitative test used here, and a diagnosis of abnormal tremor by clinicians, in a study of 63 aluminum workers.

Gerr et al. [1991] studied 79 patients referred to a clinic for neurophysiological evaluation, who had clinical examinations of vibrotactile sensitivity of index finger and great toe (scored as normal, equivocal, abnormal), quantitative tests of vibrotactile sensitivity by scoring reported sensation to a vibrating pole (the Vibraton, continuous scoring using the method of limits), and nerve conduction velocity tests of a variety of upper and lower extremity nerves. The authors found a good correlation between quantitative vibrotactile sensitivity and the clinical examination results which had three possible scores (Pearson correlation for index finger 0.66, for great toe 0.70). Our results are not directly comparable, as we had only a dichotomous classification of the clinical examination (normal/abnormal), prohibiting a meaningful calculation of a correlation coefficient. Gerr et al. did not report the predictive power of the nerve conduction tests for the clinical examination.

Gerr et al. also correlated nerve conduction velocity and amplitude for the peroneal nerve and sural nerve with the quantitative test for toe vibrotactile sensitivity (Pearson correlation coefficients −0.55 for peroneal velocity, −0.29 for peroneal amplitude, −0.28 for sural velocity, −0.29 for sural amplitude). Our corresponding Pearson correlation

coefficients were lower (−0.23, −0.13, −0.33, −0.22, respectively). It is known that nerve conduction velocity and a quantitative test of vibrotactile sensitivity do measure exactly the same neurologic phenomena [Gerr and Letz, 1993]. Nonetheless, for the lower extremity we found that both tests were good predictors of an abnormal clinical examination.

Wennemark et al. [1996] compared quantitative tests for finger vibrotactile sensitivity with self-reported symptoms and did not find a strong relationship. They did not conduct a clinical examination and thus their findings are not directly comparable to ours. However, we also found that the probability of an abnormal clinical examination for vibrotactile sensitivity was high among all top three quartiles of our quantitative test results, indicating poor predictive power for the quantitative test.

Flodmark and Lundborg [1997] studied 135 blue collar workers in manufacturing, for vibrotactile sensitivity measured quantitatively and via a clinical examination. Results are not reported in units comparable to ours, making comparison difficult. These authors do mention that abnormal clinical examinations for sensation (Phalen's test and two-point discrimination) were only increased in those with the worst finger vibrotactile sensitivity on quantitative test. In contrast we did not find that abnormal finger tactile sensitivity on clinical examination was limited to those with the worst (e.g., worst quartile) vibrotactile sensitivity on the quantitative test.

In conclusion, several quantitative tests were reasonable predictors of abnormality on clinical examination in our data. The choice of whether to rely on quantitative tests alone rather than a clinical examination in an epidemiologic field study will depend on a study's budget and goals. If the budget is limited, and if the goal of the study is to detect differences between exposed and nonexposed rather than identify subjects with clinical abnormality, then quantitative tests alone will suffice. It should be remembered that the clinical ''abnormalities'' detected in a typical clinical examination are fairly prevalent and do not generally cause any dysfunction, and that results of the clinical examination will vary neurologist to neurologist and, therefore, such results themselves are not necessarily an ideal ''gold standard.''

ACKNOWLEDGMENTS

We thank Allison Tepper, Fred Gerr, and an anonymous journal reviewer for helpful comments.

REFERENCES

Békésy G von. 1947. A new audiometer. Acta Otolaryngol (Stockh) 34:411–422.

Bhattacharya A, Shukla R, Dietrich K, Bornschein R, Berger O. 1995. Effect of early lead exposure on children's postural balance. Dev Med Child Neurol 37:861–878.

Davis-King K, Sweeney M, Wille H, Steenland K, Arezzo J. 1992. Reference values for amplitudes and conduction velocities obtained from a cohort of middle-aged and retired workers. Scand J Work Environ Health 18 (Suppl 2): 24–26.

Degowin E, Degowin R. 1981. Diagnostic examination. New York: MacMillan.

Dick R, Krieg E, Sim M, Bernard B, Taylor B. 1997. Evaluation of tremor in aluminum production workers. Neurotoxicol Teratol 19:447–453.

Dick R, Pinkerton L, Krieg E, Biagini R, Deddens J, Brightwell S, Grubb B, Taylor T, Russo J. 1999. Evaluation of postural stability in workers exposed to lead at a secondary lead smelter. NeutoToxicology 20(4):595–608.

Flodmark B, Lundborg G. 1997. Vibrotactile sense and hand symptoms in blue collar workers in a manufacturing industry. Occup Environ Med 54:880–887.

Galinsky T, Rosa R, Wheeler D. 1990. Assessing muscular fatigue with a portable tremor measurement system suitable for field use. Behav Res Methods Instrum Comput 22: 507–516.

Gerr F, Letz R, Hershman D, Farraye J, Simpson D. 1991. Comparison of vibrotactile thresholds with physical examination and electrophysiologic assessment. Muscle Nerve 14: 1058–1086.

Gerr F, Letz R. 1993. Vibrotactile threshold testing in occupational health: a review of current issues and limitations. Environ Res 60: 145–159.

Steenland K, Dick R, Howell R, Chrislip D, Hines C, Reid T, Lehman E, Laber P, Kreig E, Knott C. 2000. Neurologic function among termiticide applicators exposed to chlorpyrifos. Environ Health Perspect 108:293–300.

Wennemark M, Lundstrom R, Hagberg M, Nilsson T. 1996. Vibrotactile perception thresholds as determined by two different devices in a working population. Scand J Work Environ Health 22: 204–210.