THE SENSITIVITY AND SPECIFICITY OF TESTS FOR CARPAL TUNNEL SYNDROME VARY WITH THE COMPARISON SUBJECTS

F. GERR and R. LETZ

From the Rollins School of Public Health, Emory University, Atlanta, USA

The performance of a variety of common office-based clinical tests for detection of carpal tunnel syndrome (CTS) was assessed in 119 subjects with and without electrophysiological evidence of CTS. Symptoms compatible with CTS and electrophysiological tests positive for median mononeuropathy at the wrist were observed in 57 hands, symptoms compatible with CTS and normal electrophysiological test results were observed in 58 hands, and no symptoms compatible with CTS and normal electrophysiological test results were observed in 123 hands. For all the diagnostic tests studied, the proportion of subjects who had a false positive clinical test result was much higher in the electrophysiologically normal subjects who had CTS compatible hand symptoms than in the electrophysiologically normal subjects who were asymptomatic. These results suggest that many studies that have evaluated diagnostic tests for CTS have produced falsely optimistic estimates of the test's performance because of their use of asymptomatic comparison subjects.

Journal of Hand Surgery (British and European Volume, 1998) 23B: 2: 151–155

Because many patients with pain, tingling and numbness in the hands do not have CTS, numerous clinical tests or manoeuvres have been proposed to distinguish those with CTS as the cause of hand symptoms from those with other causes. The two most time honoured, and the two most studied, of such tests are Phalen's test (Phalen, 1966) and Tinel's sign (Dawson et al, 1990).

The usefulness of clinical tests is judged on the proportion of "positive" results obtained among those who truly have the disease (i.e. sensitivity) and the proportion of "negative" results obtained among those who are truly free of disease (i.e. specificity). A typical study to evaluate the sensitivity and specificity of any test involves identification of diseased and non-diseased subjects using an accepted best diagnostic method followed by use of the test under evaluation. The sensitivity and specificity of the test for detecting the disease in question are then readily calculated.

A review of the literature shows wide variation in the reported sensitivity and specificity of Phalen's test and Tinel's sign for detection of CTS. Estimates of the sensitivity of Phalen's test range from 10% (Golding et al, 1986) to 88% (Williams et al, 1992) and of Tinel's test from 26% (Golding et al, 1986) to 79% (Mossman and Blau, 1987). The estimates of the specificity of Phalen's test ranges from 47% (Katz et al, 1990) to 100% (Williams et al, 1992) and of Tinel's from 45% (Seror, 1987) to 100% (Williams et al, 1992). Authors' conclusions about these tests range from the statement that they "should no longer be used for diagnosis of CTS … Patients with CTS symptoms should be referred directly for neurophysiologic examination" (DeKrom et al, 1990) to "the present study demonstrates a real diagnostic and prognostic value of the wrist flexion test described by Phalen" (Seror, 1987).

There are several possible explanations for these relatively large discrepancies, including differences in the methods of performing the tests, differences in the best diagnostic test used to define the disease-negative and the disease-positive groups, and differences in selection of subjects for classification into the disease-negative and the disease-positive groups.

It has been shown in other settings that the observed sensitivity and specificity of a clinical test are, in part, dependent upon the selection of subjects for inclusion in the disease-negative and the disease-positive groups (Ransohoff and Feinstein, 1978). Studies in which investigators used asymptomatic subjects to serve as members of the CTS-free group appear to produce much higher estimates of sensitivity and specificity [e.g. Mossman and Blau (1987); Williams et al (1992)] than those studies in which the investigators used patients with hand symptoms who were found, after testing with the best available method, to be free of CTS [e.g. Katz et al (1990); DeKrom et al (1990)].

In order to test the hypothesis that use of asymptomatic subjects free of electrophysiological evidence of CTS results in improved test performance in comparison with symptomatic subjects free of electrophysiological evidence of CTS, we designed a cross-sectional study of the sensitivity and specificity of Phalen's test, Tinel's sign and several other clinical tests used for detection of carpal tunnel using two separate subgroups of subjects free of electrophysiological evidence of CTS: those without symptoms compatible with CTS (similar to CTS-negative subjects in many other published studies of these clinical tests); and those with symptoms compatible with CTS (similar to the patients on whom these tests will actually be used).

METHODS

Subjects

Symptomatic subjects were recruited from among patients referred to a university-based clinical electromyography centre. Any patient between 18 and 70 years of age with symptoms of pain, weakness, numbness or tingling in the cutaneous distribution of the median nerve that involved either hand was eligible for inclusion in the study. Asymptomatic subjects were recruited from among office and technical staff at a major university. All subjects provided informed consent approved by the Emory University Human Investigations Committee for participation in the study.

Symptoms questionnaire

A standardized questionnaire including demographics (age, gender, height and weight) and symptoms of CTS was administered to all participants. Symptoms compatible with carpal tunnel syndrome included numbness, tingling, or pain in the cutaneous distribution of the median nerve (thumb, index, long, and ring fingers) occurring during the 2 weeks before enrolment in the study. Symptoms were recorded as present only if the subject reported an intensity value of 3 or greater on a 0 ("no discomfort") to 10 ("unbearable discomfort") visual analogue scale. For the current study, subjects with symptoms compatible with carpal tunnel syndrome are referred to as "symptom-positive".

Clinical tests

All examinations were performed on study subjects bilaterally by a single board certified internist who was blinded to the results of the electrophysiological tests.

Phalen's manoeuvre was performed by maintaining maximal voluntary wrist flexion for a period of 60 seconds. The test was considered positive if there were dysaesthetic symptoms in the cutaneous distribution of the median nerve. Tinel's sign was positive when percussion over the palmar aspect of the wrist with a standard reflex hammer caused pain or paraesthesiae to radiate into the hand.

Thenar motor strength was assessed by instructing the patient to abduct the thumb while the examiner resisted the motion and pushed the thumb towards the palm. Thenar motor strength was graded as normal or abnormal. Thenar muscle mass was assessed by inspection and palpation and was graded as normal or abnormal.

Vibration perception was determined with a standard 128 Hz tuning fork. The fork was struck by the examiner and applied with moderate pressure to the bony prominence of the distal interphalangeal joint of the index fingers, bilaterally. The subject was instructed to respond verbally when the vibration was no longer perceptible. The result was recorded as normal or abnormal.

Two-point discrimination was determined on the index finger with a two-point aesthesiometer. The smallest distance perceived as two separate points was recorded in millimetres. A distance of 5 mm or more was considered abnormal.

Electrophysiological testing protocol

Standard electrophysiological testing techniques were used (Kimura, 1989). All electrophysiological measures were made with a TECA TD-20 electromyograph (TECA Corp., Pleasantville, N.Y.). Because examinations were performed for clinical indications, some variability in the details of the electrophysiological examination occurred among symptomatic individuals. Sensory and motor studies were done on the median and ulnar nerves of the affected limb in symptomatic subjects. All subjects had, at least, median motor nerve and median sensory nerve evaluation of the symptomatic hand. In asymptomatic subjects only the dominant limb was tested. Standard needle electromyography (EMG) of the abductor pollicis brevis muscle was performed on all symptomatic subjects, with the muscle at rest and during voluntary contraction. Abnormalities consistent with active or chronic denervation were recorded.

An electrophysiological study was considered positive (i.e., consistent with CTS) if any one of the following results was found:

- Median nerve distal motor latency greater than 4.4 ms at a distance of approximately 7 cm between the distal stimulation site at the wrist and the active recording electrode over the thenar eminence in a subject with normal ulnar nerve function.
- Median nerve distal motor latency of 1.8 ms or greater than the ipsilateral ulnar nerve distal motor latency.
- Median mixed nerve (sensory and motor) palm-to-wrist latency greater than 2.2 ms.
- Median sensory nerve conduction velocity from wrist-to-finger of less than 44 m/s or median sensory nerve latency from finger-to-wrist of greater than 3.8 ms.
- Isolated EMG abnormalities (active or chronic) of the abductor pollicis brevis muscle suggestive of denervation.

For the current study, subjects with electrophysiological test results compatible with carpal tunnel syndrome are referred to as "EP-positive".

Data analysis

For the purpose of determining the sensitivity and specificity of the clinical tests, it is necessary to categorize subjects as CTS-positive and CTS-negative. Consistent with clinical practice, subjects categorized as CTS-positive were those who had symptoms compatible with CTS *and* who were EP-positive. For the purpose of this paper, subjects who were EP-negative were divided into two categories: those with symptoms compatible with CTS and those without symptoms compatible with CTS. This division of EP-negative subjects was necessary to demonstrate differences in clinical test performance between subjects who differed on *symptoms* alone. The creation of a CTS-negative group with subjects who were EP-negative but who had symptoms compatible with CTS is justified by the low rate of false negative electrophysiological studies reported in the literature (Grundberg, 1983; Stevens, 1987) and the fact that clinicians do not use clinical tests for detection of CTS among patients who have no hand symptoms.

RESULTS

A total of 144 subjects were recruited. Of those recruited, 119 met criteria for inclusion in one of the three hand condition groups and were included in the analyses. Twenty-five participants had EP tests positive for a disorder other than CTS and were excluded from the analyses. Because the unit of analysis for demographic characteristics was the individual subject, and because CTS status could differ between hands in the same subject, demographic characteristics are presented for the 119 subjects stratified by dominant hand CTS status (Table 1). Demographic characteristics were virtually identical when the population was stratified by non-dominant hand CTS status and are not presented. A statistically significant difference was observed between the three groups for both mean age (F=12.73; df=2,116; $P<0.0001$) and mean body mass index (BMI) (F=5.91; df=1,116; $P=0.004$). The EP-positive group was slightly older and heavier for height than the EP-negative groups. No statistically significant differences were observed between the three groups for height and sex.

For reporting clinical test results the unit of analysis was the hand. Of the 238 hands, a total of 57 hands were categorized as CTS-positive (positive EP and symptoms compatible with CTS), and 181 hands as CTS-negative (negative EP regardless of hand symptoms, Table 2). Among the 181 EP-negative hands, 58 had CTS-compatible symptoms, and 123 did not have CTS-compatible symptoms. Of the 119 subjects, 25 had bilateral CTS (positive EP and symptoms compatible with CTS) representing 50 of the 57 CTS-positive hands. Because some subjects had unilateral CTS, the number of hands in each CTS group (Table 2) was not exactly twice the number of subjects in each CTS group when presented for the dominant hand only (Table 1).

The percentages of hands observed to have abnormal clinical test results are presented in Table 2 for each of the three groups. For each of the clinical tests studied, the percentages of abnormal results were similar between the CTS-positive group and the first of the two CTS-neg-

Table 1—Demographic characteristics of the study population stratified by dominant hand CTS group

Variable	*Total*	*CTS Positive EP* Positive Symptom positive N=30*	*CTS Negative EP* Negative Symptom positive N=30*	*CTS Negative EP* Negative Symptom negative N=59*
Age (yrs)				
Mean	42.6	50.1	43.9	38.2
SD	11.6	11.8	12.6	8.7
Height (cm)				
Mean	168	166	169	169
SD	9.6	7.4	10.3	10.2
BMI (kg/m^2)				
Mean	25.3	28.5	26.1	23.3
SD	5.7	5.6	7.1	3.9
Sex				
% Female	72	83	60	70

* EP: Electrophysiological test

Table 2—Percentage of hands with abnormal clinical test results

Outcome	*CTS Positive EP* Positive Symptom positive N=57*	*CTS Negative EP* Negative Symptom positive N=58*	*CTS Negative EP* Negative Symptom negative N=123*
Phalen's Test	43	39	3
Tinel's Sign	14	19	2
Thenar weakness	36	22	2
Thenar atrophy	16	10	0
Abnormal vibration perception (index finger)	35	17	4
Two point >= 5 mm (index finger)	16	21	1

* EP: Electrophysiological test

ative groups (i.e. EP-negative, symptom-positive hands). For example, Phalen's sign was positive among 43% of the CTS-positive hands and among 39% of the symptom-positive, EP-negative hands. Essentially this means that the clinical tests studied did not discriminate between these two groups of hands.

The percentages of abnormal clinical test results obtained in the EP-negative, symptom-negative hands (i.e. false positive results) were much lower than those obtained in the EP-negative, symptom-positive hands. For example, Phalen's test was falsely positive for 3% of the asymptomatic, EP-negative hands (specificity = 97%), whereas it was falsely positive for 39% of the symptomatic, EP-negative hands (specificity = 61%). Stated another way, the observed specificity (the percentage of disease-free hands with a negative clinical test result) of all of the tests examined was substantially better for asymptomatic hands with normal electrophysiological test results than for symptomatic hands with normal electrophysiological test results.

DISCUSSION

After accounting for the use of symptomatic versus asymptomatic disease-negative subjects, the performance of the clinical tests for CTS in the current study were consistent with those observed by others. For example, Katz et al (1990) reported a sensitivity of 32% and a specificity of 81% for "sensory loss" as a clinical test for CTS among symptomatic subjects categorized on the basis of electrophysiological evaluation. In the current study the analogous sensitivity and specificity values for loss of vibration perception were 35% and 83%. Review of other studies of these clinical tests show, consistent with the current study, that when evaluated with asymptomatic CTS-free subjects (Mossman and Blau, 1987; Williams et al, 1992), their performance was considerably better than when evaluated with symptomatic CTS-free subjects (DeKrom et al, 1990).

Our results demonstrate that the clinical tests discriminated between symptomatic and asymptomatic subjects but not between those with and without electrophysiological evidence of CTS. This indicates that the clinical tests are therefore useless in distinguishing those symptomatic patients who actually have CTS from those who do not.

This pattern of results, in which a diagnostic test performs differently among different groups of patients with the same disease status, has been observed in other settings (Lachs et al, 1992) and has been referred to as "spectrum bias" (Ransohoff and Feinstein, 1978). Indeed, the results of our study, as well as accepted epidemiologic practice, suggest that use of asymptomatic comparison subjects is not appropriate when investigating sensitivity and specificity of clinical tests. The true utility of a clinical test is not whether it can distinguish those for whom the diagnosis would never be entertained from those who have a clinical presentation that is typical for the disorder. Instead, diagnostic tests are used to assist in identifying those with the condition from among a population of patients with symptoms compatible with the condition, some of whom actually have it and some of whom do not.

While performing the clinical examination, the examiner was blinded to the CTS group status (symptoms and electrophysiological test results) of all study subjects. Furthermore, the examiner, board certified in Internal Medicine and Occupational Medicine, was highly qualified to perform these evaluations. Therefore, any potential errors in the outcomes of the clinical tests would be expected to be small and evenly distributed across all the CTS groups. Such a pattern of error is not expected to result in the systematic differences observed in the current study.

Another possible criticism of the current study was the use of electrophysiological studies to categorize subjects into CTS-positive and CTS-negative groups for determination of disease status. Some authors have suggested that electrophysiological studies are not sufficiently sensitive for detection of CTS and that, therefore, CTS can be present despite normal electrophysiological studies. For example, Grundberg (1983) has suggested that the false negative rate of electrodiagnostic evaluation for detection of CTS is as high as 8%. Such false negatives would affect only one of the three groups used in the current study, the EP-negative, symptom-positive group which was considered, in the current study, to be free of CTS. Even assuming that such false negatives were present in the current study, the resultant increase in the proportion of abnormalities would be only a fraction of the proportion of abnormality actually observed. For example, if as many as half of the EP-negative, symptom-positive subjects actually had CTS, and Phalen's sign were 97% specific (the rate from the EP-negative, symptom-negative group) and 43% sensitive (the rate from the EP-positive, symptom-positive group) then the proportion of positive results in this EP-negative, symptom-positive group would be 23% rather than the observed 39%. For this reason, the differences in test results between the two CTS-negative groups cannot be attributed solely to subjects in the symptomatic group who, despite normal electrophysiology, might have actually had CTS.

Our results demonstrate that spectrum bias affects the observed performance of commonly used clinical tests for CTS and is likely to be responsible for the widely divergent estimates of sensitivity and specificity for Phalen's test and Tinel's sign observed in the published literature. In addition, considerable doubt is cast on the conclusions of studies of new diagnostic tests for identification of carpal tunnel syndrome because of falsely elevated estimates of sensitivity and specificity that can result if asymptomatic subjects constitute the disease free comparison group.

Acknowledgements

This work was supported by a grant from the National Institute for Occupational Safety and Health (K01-OH00098). The authors wish to thank Dr Linton C. Hopkins, Director of the Emory Electromyography Center, for his support of this project and Ms Deborah Harris-Abbott for her assistance with programming and data management. The authors also wish to thank an anonymous reviewer who made numerous helpful suggestions.

References

Dawson DM, Hallett M, Millender LH. *Entrapment neuropathies,* 2nd edn. Boston Little, Brown and Co., 1990: 25.

DeKrom MCTFM, Knipschild PG, Kester ADM, Spaans F (1990). Efficacy of provocative tests for diagnosis of carpal tunnel syndrome. The Lancet, 335: 393–395.

Golding DN, Rose DM, Selvarajah K (1986). Clinical tests for carpal tunnel syndrome: An evaluation. British Journal of Rheumatology, 25: 388–390.

Grundberg AB (1983). Carpal tunnel decompression in spite of normal electromyography. Journal of Hand Surgery, 8A: 348–349.

Katz JN, Larson MG, Sabra A. et al (1990). The carpal tunnel syndrome. Diagnostic utility of the history and physical examination findings. Annals of Internal Medicine, 112: 321–327.

Kimura J. *Electrodiagnosis in diseases of nerve and muscle: Principles and practice*, 2nd edn. Philadelphia: F.A. Davis Co., 1989: 104–109.

Lachs MS, Nachamkin I, Edelstein PH, Goldman J, Feinstein AR, Schwartz JS (1992). Spectrum bias in the evaluation of diagnostic tests: Lessons from the rapid dipstick test for urinary tract infection. Annals of Internal Medicine, 117: 135–140.

Mossman SS, Blau JN (1987). Tinel's sign and the carpal tunnel syndrome. British Medical Journal, 294: 680.

Phalen GS (1966). The carpal-tunnel syndrome. Journal of Bone and Joint Surgery, 48A: 211–228.

Ransohoff DF, Feinstein AR (1978). Problems of spectrum bias in evaluating the efficacy of diagnostic tests. New England Journal of Medicine, 299: 926–930.

Seror P (1987). Tinel's sign in the diagnosis of carpal tunnel syndrome. Journal of Hand Surgery, 12B: 364–365.

Stevens JC (1987). AAEE Minimonograph #26: The electrodiagnosis of carpal tunnel syndrome. Muscle & Nerve, 10: 99–113.

Williams TM, Mackinnon SE, Novak CB, McCabe S, Kelley L (1992). Verification of the pressure provocative test in carpal tunnel syndrome. Annals of Plastic Surgery, 29: 8–11.

Received: 9 January 1997
Accepted after revision: 1 September 1997
F. Gerr MD, Department of Environmental and Occupational Medicine, Rollins School of Public Health at Emory University, 1518 Clifton Road, Atlanta, GA 30322, USA.