

CRITICAL APPRAISAL OF THE LITERATURE ON THE NC-STAT DEVICE FOR NERVE CONDUCTION STUDIES

Beverly C. Walters, M.D., M.Sc., FRCSC, FACS

EBMAdvisors



BEVERLY C. WALTERS, M.D., M.Sc., FRCSC, FACS Biography

Dr. Walters is a neurosurgeon and clinical epidemiologist who is retired from a full time clinical practice in general neurosurgery, focusing on spine, brachial plexus, and peripheral nerve, to pursue her interest in applied evidence-based medicine. Until recently, she maintained her academic appointments as Clinical Professor of Neurological Surgery at New York University Medical School, Adjunct Professor of Neurosurgery of Surgery at the Uniformed Services University of the Health Sciences in Bethesda, and Visiting Professor of Neurosurgery at the University of Zagreb, Croatia. She was also Director of Outcomes Research at Walter Reed Army Medical Center. Her experience in clinical research involves the design and implementation of randomized controlled trials, case-control studies, a meta-analysis, and she has written extensively on study design and assessment of the quality of medical literature, including being the author/co-author of fourteen monographs on critical appraisal, as well as editorials and book chapters on statistical aspects of neurosurgical research. She has extensive experience in survey methodology and has developed a statistical technique for use in generating evidence for use in algorithms when randomized controlled trials cannot be done. She has been a lecturer, tutor, and course designer for several educational programs on critical appraisal of the medical literature for practicing clinicians, graduate students, postgraduate trainees, and medical students. She is the co-editor of the neurosurgical textbook, Evidence-based Neurosurgery, and has a consulting practice in evidence-based medicine under the name of EBM Advisors.

For many years, Dr. Walters was the chairperson of the Practice Guidelines Committee of the American Association of Neurological Surgeons, a member of the Outcomes Committee of the AANS/CNS, and a member of the Committee for the Assessment of Quality of the AANS/CNS. Her guidelines activities have included the development of the Guidelines for the Management of Severe Head Injury, the Guidelines for Prehospital Management of Traumatic Brain Injury, Guidelines for Penetrating Brain Injury, Spinal Cord Injury Guidelines, Guidelines for the Treatment of Low Grade Gliomas, and the Guidelines for the Management of the Neurobehavioral Consequences of Traumatic Brain Injury. Her work on the Spinal Cord Injury Guidelines was recognized by Congress for its achievement. She gives courses on guideline development, and writes frequently on outcome measurement for neurosurgeons.

CRITICAL APPRAISAL OF THE LITERATURE ON THE NC-STAT DEVICE FOR NERVE CONDUCTION STUDIES

Introduction

The purpose of this document is to evaluate the utility of an automated nerve conduction study instrument (NC-stat) provided at the point of care. This evaluation is based upon a review of the literature on the NC-stat and the reference comparative study, the “traditional” nerve conduction study (NCS). Prior to undertaking such a review, it is important to understand the underlying assumptions for the task at hand. When evaluating diagnostic tests, it is essential to adopt the appropriate orientation to examining the literature. Unfortunately, many physicians have very little training in the different sorts of literature encountered in evaluating patient management. Because of the focus upon efficacy and effectiveness in clinical trials of preventive and therapeutic interventions, the same orientation is sometimes brought to the critical appraisal of the literature on diagnostic tests, quite erroneously. Whereas treatment effectiveness is tested using some high-quality comparative study design (randomized controlled trial, case-control study) in which patient outcome is the key measure of success, diagnostic tests are judged by their accuracy, reliability and validity. Validity is the extent to which the test outcomes match some “gold standard” reference test representing true disease presence or absence. Accuracy is represented by sensitivity and specificity, positive and negative predictive values, and likelihood ratios. Reliability is a measure of the test’s precision, or the extent to which the test results are reported similarly time after time. Therefore, studies of diagnostic accuracy and precision require quite different evaluation criteria than therapeutic effectiveness.

In the case of nerve conduction studies, there is no true “gold standard”, because there are no studies that have compared any procedure for obtaining NCS to *in vivo* studies carried out directly upon exposed nerves in either animals or humans. Therefore, traditional nerve conduction studies function as an accepted “reference test” to which an “index test” such as NCS using the NC-stat are compared for diagnostic accuracy. This is unfortunate, because in some situations the index test may be superior to the reference test, especially when they are aimed at the same process, i.e., measuring velocity, latency, amplitude, waveform morphology, etc., of electrical nerve function. For example, in a study by Wells et al (1), the NC-stat device was better able to predict lumbar radiculopathy as found on MRI, clinical history and physical examination than had been seen previously in other evaluations of the predictive value of traditional nerve conduction studies. This particular observation was further explored and explained using mathematical modeling and additional analysis of the data from Wells et al. by Gozani, Kong & Fisher (2). However, if the reference test is widely accepted, as traditional nerve conduction is, then what must be looked at closely is the concordance of the tests in given diagnostic situations. For its review of the NC-stat device for consideration of granting 510(k) clearance for its use, the FDA considered the devices essentially the same, and the NC-stat received this clearance. What will be outlined below is an evaluation of the peer-reviewed literature on the NC-stat device, and conclusions regarding its place in the diagnostic armamentarium of the physician dealing with patients with peripheral nervous system disorders.

Device Description

The NC-stat system is comprised of three interrelated modules or “steps”. These include the biosensors, the instrument, and the onCall reporting system. The biosensors include an array of stimulating, recording, and reference electrodes, microelectronic components, and electrical circuits that are configured for each specific nerve requiring evaluation. These are placed using anatomic landmarks on the hand, wrists, elbow, ankles, and foot. The instrument is a digital device that acquires the waveforms, and performs software-guided waveform analysis, similar to traditional NCS, as well as providing comparative normative data visually. The onCall reporting system includes a docking station for the instrument with a high-speed modem that contacts the manufacturer for comparison to a large database of normal and diseased patient data and provides a report analyzing the information in detail. The study itself can be carried out by someone trained in the physician’s office staff or the physician him/herself, whereas the report is interpreted solely by the physician in the context of the patient’s history and physical examination, and other clinical data. (3)

Diagnostic Accuracy of the NC-stat Device

Several articles evaluating the NC-stat device as a diagnostic test have been published, beginning in 2000, the most recent being published in 2008. (4 - 11) These studies have evaluated the device for use in carpal tunnel syndrome, ulnar neuropathy, lumbosacral radiculopathy, and diabetic peripheral neuropathy. In order to evaluate the quality of the studies with respect to evaluating diagnostic accuracy, the Standards for Reporting of Diagnostic Accuracy (STARD) developed by scientists and editors of peer-reviewed scientific journals (published simultaneously in the *Annals of Internal Medicine*, *Clinical Chemistry*, *Radiology*, *BMJ*, *The Lancet*, *American Journal of Clinical Pathology*, *Clinical Biochemistry*, *Clinical Chemistry and Laboratory Medicine*, and *Journal of Clinical Microbiology*) can be utilized. (12) There are 25 key items in a checklist to be applied to articles on diagnostic accuracy to ascertain their validity, according to these standards. They are shown in Table 1. Not all of these criteria are of equal importance, nor are all of them applicable to nerve conduction studies, where both the traditional, or reference, test and the index test (NC-stat) are performed in the same manner, electrophysiologically speaking. These provide an excellent guide for the authors of articles reporting diagnostic test accuracy when writing their manuscripts, and for reviewers of journals. For the purposes of this evaluation, a different, more succinct, and practical method for evaluating diagnostic tests can be used. (13) Several simple questions can be posed (and are the key questions upon which the STARD guidelines were based) to quickly evaluate diagnostic test accuracy. They include three main questions with their subordinate questions:

1. Is this evidence about a diagnostic test valid?
 - a. Was there an independent, blind comparison with a reference standard of diagnosis?
 - b. Was the diagnostic test evaluated in an appropriate spectrum of patients?
 - c. Was the reference standard applied regardless of the test results?

- d. Was the test evaluated in a second, independent group of patients?
2. Does this (valid) evidence demonstrate an important ability of this test to accurately distinguish patients who do and do not have a specific disorder?
 - a. Sensitivity and specificity calculated, or able to be calculated?
 - b. Positive and negative predictive values calculated, or able to be calculated?
 - c. Other tests of accuracy (e.g., likelihood ratios, ROC curves)?
3. Can this valid, important test be applied to a specific patient?
 - a. Is the diagnostic test available, accurate, affordable, and precise in the target setting?
 - b. Will the test results affect clinical management and patient outcome?

Each of the papers found by Medline search that evaluated the NC-stat device was submitted to appraisal of clinical utility and diagnostic accuracy using subsets of the first two key questions regarding diagnostic accuracy. The blinded comparison is assumed, because the device uses automated data acquisition and an automated comparison to a large registry of normal and abnormal data of various kinds, eliminating variability and assuring reliability. Multiple types of patients have been evaluated in several different studies, satisfying the requirement for validation in at least a second group of patients. Therefore, the most important issues to examine are whether there is an appropriate spectrum of patients in the study, whether the reference test is administered in all the studied subjects, and whether diagnostic accuracy (acceptable sensitivity, specificity, correlation, ROC curves, etc.) has been determined. The results of this critical appraisal of the literature are found in Table 2. According to this evaluation, the remaining key question is answered: i.e., the NC-stat device provides reliable and accurate assessment of peripheral nerve function compared to traditional nerve conduction studies, and is readily available and affordable. Like traditional nerve conduction studies, clinical management is affected by evaluation of peripheral nerve function, and therefore the potential to affect patient outcome is also affected, assuring clinical utility. This has been shown in a study by Megerian et al (14), in which Family Medicine physicians, Primary Care providers, and Internal Medicine physicians were able to successfully use the NC-stat in their practices for screening of patients for carpal tunnel syndrome, meeting evidence-based guidelines for using diagnostic tests. Further, in a study by Kong et al (15), clinical utility of the device in diagnosing diabetic peripheral neuropathy in a wide spectrum of patients in various practices – Primary Care, Internal Medicine, Endocrinology, Podiatry, Rheumatology, Orthopedic Surgery, Occupational Medicine, Neurology, Pain Management – using over 63,000 diagnostic tests was shown.

Research Using the NC-stat Device for Clinical Measurement

Because of its reliability and accuracy, the device has been used for establishing baseline and outcome data in several important studies where it was the only device used for measuring nerve

conduction. Prognostic factors producing mononeuropathy of the upper extremity at the wrist among engineers who used computers continually in their work were studied by Conlon and Rempel (16) at the University of California San Francisco School of Public Health. Their Methods and Materials section describes their nerve conduction testing as follows:

Nerve Testing

Upper extremity peripheral nerve function at the wrist was tested in each participant using an automated electrophysiological neurodiagnostic device (NC-stat; NeuroMetrix, Inc.). The device measures distal motor latency (DML) and has been found to have good accuracy when compared with conventional methods. Four tests were performed using a different biosensor for the ulnar nerve and the median nerve in both extremities. The biosensor consists of two stimulation electrodes, two detection electrodes, a temperature electrode, and one reference electrode. Biosensors for the median nerve come in small, medium, and large and are selected based on the participant's weight. Motor latency is measured by stimulating the nerve proximal to the wrist. The electrical stimuli evoke myoelectric responses in the innervated muscles. The biosensor detects these myoelectric responses, which are compound muscle action potentials, as electrical potentials on the skin. For example, with the median nerve test, the stimulus cathode is located 3 cm proximal to the distal wrist crease. The compound muscle action potential of the abductor pollicis brevis (APB) is detected through volume conduction by electrodes located proximal to the wrist crease. The measurements used in this study included the DML for the ulnar and median nerves at the wrist. The device calculates the median or ulnar distal motor latency (calculated DMLs) and adjusts for skin temperature. The collected responses to stimuli are downloaded to the NeuroMetrix server and normative values, adjusted for age, gender, and height, are returned as temperature-corrected DML values and as percentiles. In the logistic analyses, an abnormal DML was defined as a value at or below the population fifth percentile. Our case definition of ulnar or median entrapment neuropathy required the presence of an abnormal DML and the participant's rating of ipsilateral distal upper extremity discomfort greater than 1 on the scale of 0 to 10.

This use of the device in clinical research attests to the belief of the scientists in the reliability and validity of the device, as well as its ease of administration. More recently, these authors went on to design and carry out a randomized controlled trial of interventions aimed at reducing the upper extremity disorders they had found among the previously-studied population of engineers, using the same methods and materials for assessing baseline and outcome data with NC-stat. (17)

In an earlier study supported by the American Foundation for Surgery of the Hand, Guyette and Wilgis (18) from the Curtis National Hand Center in Baltimore investigated the factors predicting success with surgical treatment for carpal tunnel syndrome and the time course for observation of improvement. The electrophysiological studies used for baseline and outcome in their investigation were produced by the NC-stat device. In the same year, Mani et al (19) at Albert Einstein College of Medicine investigated the effective dose of chemotherapy on solid tumors in a Phase I trial. One of the unwanted side effects of the treatment is neurosensory deficit with peripheral neuropathy. In order to measure the presence or absence of this side effect with the proposed new chemotherapeutic agent, nerve conduction studies were carried out. Their Patients and Methods section describes these measurements as follows, indicating their belief in the appropriateness of the outcome measurement:

***Neurological Assessments.** Vibration perception threshold (VPT), distal motor latency (DML), and F-wave latency values were obtained at baseline and at the end of every two courses of therapy until disease progression or onset of intolerable toxicity. The same measurements were also obtained at 8 weeks after the discontinuation of drug.... DML is the interval between the stimulus and the onset of the compound*

muscle action potential in the thenar muscles. F-wave latency is the median interval between the stimulus and the onset of an action potential in the thenar muscle resulting from antidromic activation of motor neurons in the spinal cord. These parameters were measured in the median nerve using NC-stat (Neurometrix, Inc., Boston, MA), which has a 90% sensitivity and specificity in detecting median nerve entrapment or systemic neuropathies. NC-stat is noninvasive and requires ~6 min to set up and use. Detailed neurological assessments using these techniques have been reported in other studies of peripheral neuropathy.

In a study of the incidence of diabetic peripheral neuropathy in primary care settings, Vinik et al (20) performed NCS in 1,434 patients at 28 centers, and identified several factors predictive of DPN among the study participants. Their Materials and Methods section describes very elaborate and sophisticated measurements of nerve conduction:

NCS for identification of DPN and determination of severity

Subjects in both cohorts received an automated NCS (NC-stat®, NeuroMetrix, Inc., Waltham, MA) that included a bilateral evaluation of the deep peroneal nerve. Each site identified allied health professionals on staff to perform the NCS. The operator was trained according to the manufacturer's instructions for use. The raw NCS data and automated decision support analysis were made available to the physician for use in diagnosis and patient management. The parameters recorded on bilateral deep peroneal nerve stimulation included the distal motor latency (DML), compound muscle action potential amplitude (CMAP) measured from baseline to negative peak, the mean F-wave latency (FWL) among at least three and up to 20 responses (or 40 stimuli), and the presence of one or more A-waves. The dependence of these parameters (except A-wave) on subject age, height, weight, body mass index (BMI), and skin surface temperature was examined in the Control cohort using multiple linear regression. Age, height, and skin surface temperature were related to DML and FWL ($P < 0.05$), age was related to CMAP ($P < 0.05$), and weight and BMI were not correlated to any parameter ($P < 0.05$). The DML and FWL were normalized to a standard temperature of 30°C, age 40 years, and height 172 cm. The CMAP was normalized to age 40 years. No A-wave adjustments were made. The temperature correction factors were 0.18 ms/°C and 0.38 ms/°C for the DML and FWL, respectively. The respective age and height correction factors were 0.006 ms/year and 0.02 ms/cm for the DML and 0.05 ms/year and 0.30 ms/cm for the mean F-wave latency. CMAP correction was performed on the square root of the amplitude with an age correction factor of < 0.005 $\mu\text{mV}/\text{year}$.

Hardy et al (21), in an industry-sponsored metanalysis of three randomized controlled trials of treatment for diabetes in patients with diabetic neuropathy published in *Diabetes Care*, reported that two of the three trials used the NC-stat device for measurement of diabetic peripheral neuropathy. Their Research Design and Methods section states: “Nerve conduction studies were performed in two of three studies at baseline and at the conclusion of both the acute and extension phases using the NCstat automated nerve conduction testing system (NeuroMetrix, Waltham, MA). Parameters measured included peroneal A wave (presence or absence), deep peroneal F wave latency, ulnar F wave latency, ulnar distal sensory latency, and deep peroneal compound action potential amplitude.” This once again demonstrates the acceptability of the NC-stat device as an adequate measure of peripheral neuropathy in clinical research.

Timpson et al (22) performed a study to examine the intraoperative effects of tourniquet application and local anesthetic injection on nerve monitoring during carpal tunnel surgery. The intraoperative studies were done using the NC-stat device. Their Methods section describes the intraoperative measurements as follows:

Stimulation and Recording Procedures

Electrodiagnostic monitoring was performed with a nerve conduction testing system (NC-stat®; NeuroMetrix, Inc., Waltham, Mass) using prefabricated median motor/sensory biosensors. At the beginning of each study, a trained technician placed a biosensor, and it remained in place throughout the investigation. SNAPs were recorded from the third digit after median nerve stimulation 6 cm proximal to the distal wrist crease, leaving the field open for lidocaine administration. Before intervention, a test was performed (M0) to determine the supramaximal stimulation level for each nerve by applying a stimulus ramp sequence from 10 to 60 mA at 200 µs and step size of 5 mA. Supramaximal stimulation was defined as the stimulus level at which the peak-to-peak amplitude variation for 3 consecutive SNAPs was no higher than 5% from the mean of the 3 SNAP amplitudes. All subsequent nerve conduction tests (M1-M8) recorded 10 SNAPs using the previously determined supramaximal level. Immediately before intervention, a baseline measurement was taken (M1). Beginning 1 minute after intervention, the nerve conduction testing system was used at 2-minute intervals throughout the 15-minute investigation. Surface temperature measurements were made at the stimulation site. All tests were stored in the nerve conduction testing system and downloaded to a computer for analysis at the end of each investigation. The 10 individual waveforms were averaged to form the mean SNAP for each test. Assessing the feasibility of intraoperative monitoring involved calculating peak-to-peak amplitude change as a percentage of the baseline measurement (M1).

This sample of clinical trials recognizing the validity of the NC-stat device as a measure of peripheral nerve function similar to traditional, more difficult to obtain nerve conduction studies further underscores the clinical utility of the device in clinical practice.

Conclusions

The NC-stat device has been studied assiduously and represents a further development of automation of nerve conduction studies with acceptable diagnostic accuracy when compared to traditional nerve conduction studies for which automation has also become an industry standard. Like traditional studies, the physician must still use his or her clinical acumen in interpreting the test in the context of the patient requiring management, including diagnosis and treatment. This evaluation has not uncovered any potential for invalidity of the device or its ability to perform the desired electrophysiological evaluation of the patient with suspected or potential peripheral nerve disorders. Because it has been shown to be straightforward to use at the point of service, it has the potential to provide availability of nerve conduction studies that might otherwise be prohibitive in terms of less available, more expensive testing.

REFERENCES

1. Wells MD, Meyer AP, Emley M, Kong X, Sanchez R, Gozani S. Detection of Lumbosacral Nerve Root Compression with a Novel Composite Nerve Conduction Measurement. 2002 SPINE 27, (24), 2811–2819
2. Gozani SN, Kong X, Fisher MA. Factors influencing F-wave latency detection of lumbosacral root lesions using a detection theory based model. 2006 Clinical Neurophysiology 117: 1449–1457
3. Gozani SN, Fisher MA, Kong X, Magerian JT, Rutcove SB. Electrodiagnostic Automation: Principles and Practice. 2005 Phys Med Rehabil Clin N Am 16: 1015-1032
4. Leffler CT, Gozani SN, Nguyen ZQ, Cros D. An Automated Electrodiagnostic Technique for Detection of Carpal Tunnel Syndrome. 2000 Neurology and Clinical Neurophysiology 2000 (3A), 2-10
5. Rotman MB, Enkvetchakul BV, Megerian JT, Gozani SN. Time Course and Predictors of Median Nerve Conduction After Carpal Tunnel Release. 2004 J Hand Surg 29A:367–372
6. Elkowitz SJ, Dubin NH, Richards BE, Wilgis EF. Clinical utility of portable versus traditional electrodiagnostic testing for diagnosing, evaluating, and treating carpal tunnel syndrome. 2005 Am J Orthop. Aug;34 (8):362-4.
7. Perkins BA, Grewal J, Ng E, Ngo M, Bril V. Validation of a Novel Point-of-Care Nerve Conduction Device for the Detection of Diabetic Sensorimotor Polyneuropathy. 2006 Diabetes Care 29:2023–2027
8. Jabre JF, Salzsieder BT, Gnemi KE. Criterion validity of the NC-stat automated nerve conduction measurement instrument. 2007 Physiol Meas 28: 95–104
9. Armstrong TN, Dale AM, Al-Lozi MT, Franzblau A, Evanoff BA. Median and Ulnar Nerve Conduction Studies at the Wrist: Criterion Validity of the NC-stat Automated Device. 2008 J. Occup Environ Med. Accepted for publication.
10. Fisher MA, Bajwa R, Somashekar KN. Routine electrodiagnosis and a multiparameter technique in lumbosacral radiculopathies. 2008 Acta Neurol Scand : 10: 1-4
11. Perkins BA, Orszag A, Grewal J, Ng E, Ngo M, Bril V. Multi-Site Testing With a Point-of-Care Nerve Conduction Device Can Be Used in an Algorithm to Diagnose Diabetic Sensorimotor Polyneuropathy. 2008 Diabetes Care 31:522–524
12. Bassuyt PM, Reltsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HCW, Lijmer JG. The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration. 2003 Ann Intern Med 138: W1-W12
13. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence Based Medicine: How to Practice and Teach EBM. 2000 Churchill Livingstone, Toronto, pp 67-93

14. Megerian JT, Kong X, Gozani SN. Utility of Nerve Conduction Studies for Carpal Tunnel Syndrome by Family Medicine, Primary Care, and Internal Medicine Physicians. 2007 J Am Board Fam Med 20:60–64
15. Kong X, Lesser EA, Potts FA, Gozani SN. Utilization of Nerve Conduction Studies for the Diagnosis of Polyneuropathy in Patients with Diabetes: A Retrospective Analysis of a Large Patient Series. 2008 Journal of Diabetes Science and Technology 2, (2): 268-274
16. Conlon CF, Rempel DM. Upper Extremity Mononeuropathy Among Engineers. 2005 J Occup Environ Med. 47:1276–1284
17. Conlon CF, Kraus, Rempel DM. A randomized controlled trial evaluating an alternative mouse and forearm support on upper body discomfort and musculoskeletal disorders among engineers. 2008 Occup Environ Med 65(5):311-318
18. Guyette TM, Wilgis EFS. Timing of Improvement After Carpal Tunnel Release. 2004 Journal of Surgical Orthopaedic Advances 13(4):206–209
19. Mani S, McDaid H, Hamilton A, Hochster H, Cohen MB, Khabelle D, Griffen T, Lebwohl DE, Liebes L, Muggia F, Horwitz SB. Phase I Clinical and Pharmacokinetic Study of BMS-247550, a Novel Derivative of Epothilone B, in Solid Tumors. 2004 Clin Cancer Res 10: 1289-1298
20. Vinik AI, Kong X, Megerian JT, Gozani SN. Diabetic Nerve Conduction Abnormalities in the Primary Care Setting. 2006 Diabetes Technology and Therapeutics 6 (8): 654 – 662
21. Hardy T, Sachson R, Shen S, Armbruster M, Boulton AJM. Does Treatment With Duloxetine for Neuropathic Pain Impact Glycemic Control? 2007 Diabetes Care 30: 21-26
22. Timpson WL, Kong X, Hamlet WP, Gross P, Gozani S. Time-Dependent Changes in Median Nerve Sensory Amplitude After Local Anesthetic Administration and Tourniquet Application. 2006 Am J Orthop 35(11): 515–519.

Table 1: Standards for Reporting of Diagnostic Accuracy (STARD)

- 1. Identify the article as a study of diagnostic accuracy.**
- 2. State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.**
- 3. Describe the study population: The inclusion and exclusion criteria, setting and locations where data were collected.**
- 4. Describe participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?**
- 5. Describe participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected.**
- 6. Describe data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?**
- 7. Describe the reference standard and its rationale.**
- 8. Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.**
- 9. Describe definition of and rationale for the units, cutoffs, and/or categories of the results of the index tests and the reference standard.**
- 10. Describe the number, training, and expertise of the persons executing and reading the index tests and the reference standard.**
- 11. Describe whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.**
- 12. Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95% Confidence Intervals).**
- 13. Describe methods for calculating test reproducibility, if done.**
- 14. Report when study was done, including beginning and ending dates of recruitment.**
- 15. Report clinical and demographic characteristics of the study population (e.g., age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).**

- 16. Report the number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).**
- 17. Report time interval from the index tests to the reference standard, and any treatment administered between them.**
- 18. Report distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.**
- 19. Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results report the distribution of the test results by the results of the reference standard.**
- 20. Report any adverse events from performing the index tests or the reference standard.**
- 21. Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g., 95% Confidence Intervals).**
- 22. Report how indeterminate results, missing responses, and outliers of the index tests were handled.**
- 23. Report estimates of variability of diagnostic accuracy between subgroups of participants, readers, or centers, if done.**
- 24. Report estimates of test reproducibility, if done.**
- 25. Discuss the clinical applicability of the study findings.**

Table 2. Critical Appraisal of Articles on Diagnostic Accuracy of NC-stat

Citation	Appropriate Spectrum of Patients?	Reference Standard Consistently Applied?	Diagnostic Accuracy Calculated?
Leffler et al 2000 ⁴	75 patients with upper extremity or neck symptoms, ages 18-75 years	Yes. Although only to symptomatic subjects, asymptomatic hands were used as controls.	Correlation .90 (p<.001), Sensitivity (87%), specificity (90%), ROC curves (Odds ratio 8.81, 95% C.I. - range, 4.04 – 19.21)
Rotman et al 2004 ⁵	48 patients with carpal tunnel syndrome, ages 21-74 years	Yes. All patients had both studies performed.	Pearson Correlation Coefficient = .94, p<.001
Elkowitz et al 2005 ⁶	72 patients with carpal tunnel syndrome, ages 30-78 years	Yes. All patients had both studies performed.	Linear regression analysis carried out (r=.88), p<.001
Perkins et al 2006 ⁷	72 patients with diabetes	Yes. All patients had both studies performed.	Spearman Correlation Coefficient = .95, p<.0001, Sensitivity (92%), specificity (82%), positive predictive value (92%), negative predictive value (82%), overall accuracy (89%)
Jabre et al 2007 ⁸	60 patients with lower extremity complaints, ages 17–83 years	Yes. All patients had both studies performed.	Spearman Correlation Coefficient: Peroneal nerve: .70 (DML), .86 (AMP), .91 (FLAT), Posterior tibial nerve: .45 (DML), .73 (AMP), .90 (FLAT). All correlations were statistically significant (p<0.01 or better)
Armstrong 2008 ⁹	33 patients referred for upper extremity neuropathy	Yes. All patients had both studies performed.	Pearson correlation coefficients ranged from .39, p=.03 (median CMAP) to .94, p<.001 (median and ulnar SNAP); ROC curve ranges = .91 (95% CI = .78 – 1.00) - .98 (95% CI = .94 – 1.00)
Fisher 2008 ¹⁰	34 patients with lumbosacral radiculopathy	Yes. All patients had both studies performed.	Sensitivity and specificity are reported for each nerve and each aspect of the test and range from .50-.92 (sensitivity) and .40-1.00 (specificity); raw agreement ranged from .58 to .79
Perkins 2008 ¹¹	72 patients with diabetic peripheral neuropathy	Yes. All patients had both studies performed.	Spearman correlation coefficients ranged from .76 (median motor F-wave latency) to .91 (sural sensory amplitude), All correlation coefficients significant at p<.0001