An Evidence Based Technology Assessment of the NC-stat® Device

Elizabeth Brown, MD

August 21, 2008

Table of Contents	**Page**

Executive Summary	2
Introduction	4
Methods	8
Test-Retest of NC-stat	9
Comparative Studies of NC-stat and Traditional Methods	9
Discussion	12
Summary	18
Glossary of Terms	19
References	21
Elizabeth Brown, MD. Biography.	24

Executive Summary

Devices used for nerve conduction studies (NCS) have steadily evolved over the decades, such that all commonly used devices incorporate some level of automation in both data collection and interpretation. The NC-stat® device represents a highly automated version of what is considered a "traditional" NCS device. However, since the degree of automation represents a continuum, it is difficult to define specific characteristics of a "traditional" device. The NC-stat device is designed to facilitate standardized electrode placement and to automate the NCS testing up to physician interpretation/diagnosis. This high degree of automation suggests that a broader range of physician specialties may offer NCS studies in different settings, such as at the bedside or in a primary care setting. While some may have questioned the ability of non-neurologists to interpret NCS studies, this is an issue of credentialing that is not specific to the NC-stat device, but should apply to any NCS study, regardless of the underlying technology.

Validation of the NC-stat technology is initially based on its technical performance. Novel diagnostic technologies are typically further assessed based on their diagnostic performance and the impact of the diagnosis on the management of the patient. However, the NC-stat device does not represent a novel diagnostic technology, but rather a further evolution of existing NCS methods. The diagnostic performance of NCS studies is an accepted component of the assessment of neuropathies with clear implications for patient management. Therefore, these downstream health outcomes associated with the NC-stat device can be extrapolated from traditional NCS methods.

Therefore, the technical performance of the NC-stat devices emerges as the key assessment parameter. The technical performance was assessed based on test-retest studies and comparative studies with traditional devices as reported in prospective case series. Key statistics include various correlation measures, such as intraclass coefficients, Spearman or Pearson correlation coefficients. All reviewed studies reported high correlations for the parameters studied with few exceptions, which is not surprising since both NC-stat and "traditional" NCS devices are based on the same underlying physiology. However, it is important to realize that comparisons to traditional NCS studies are limited by the inherent variability in any NCS study, due to the multistep nature of the procedure in which minor variations can result in varying results in one or more of the multiple different parameters studied. Therefore, it is relevant to compare the results of NC-stat with the results of test-retest studies of traditional NCS.

Two large studies of traditional methods have been undertaken to determine the test-retest reliability of NCS as an outcome measure for different therapies for diabetic neuropathy. Results were presented as the coefficient of variation (COV), where an increasing value reflects increasing variability. The COVs ranged from 3-53%, depending on the nerve and parameter studied. In addition, from 19-34% of studies did not meet specified data quality criteria developed by a central laboratory. These studies illustrate that "traditional" NCS studies as performed by multiple laboratories are a flawed gold standard. Therefore, the reported correlation coefficients of NC-stat vs. traditional methods reflect the limitations of the traditional methods as well as the comparative performance of an NC-stat study. Essentially, correlation studies between NC-stat and traditional NCS can only be reasonably expected to demonstrate up to the level of reliability shown in test-retest studies of traditional methods.

Data from studies comparing the technical performance of NC-stat and traditional methods show excellent correlation that is similar in pattern and range for test-retest studies using traditional NCS methods. These results support the use of the NC-stat device as an alternative to traditional NCS methods.

Introduction

For the past 60 years, nerve conduction studies (NCS) have evolved to a standard and essential technique to diagnose diseases of the peripheral nervous system. NCS assess the speed (conduction velocity and/or latency), size (amplitude) and shape of the response of motor, sensory or mixed nerves. Late responses, such as F-wave and H-reflex studies, may also be recorded. Motor NCS are performed by applying electrical stimulation at various points along the course of a motor nerve and recording the electrical response from the corresponding muscle. Sensory nerve conduction studies are performed by applying electrical stimulation over a sensory nerve and recording at another site along the nerve.[1]

Similar to many technologies, the sophistication of NCS studies has evolved over the years, incorporating many automated features in data acquisition, data analysis and data reporting. In general, automation is designed to facilitate the study performance, increase reproducibility and assist in interpretation. For example, the earliest NCS studies were performed with analog equipment; physicians were required to manually set all the parameters of the study, such as setting the sweep speed of the oscilloscope and setting the filter setting to optimize data. Data analysis was performed manually based on paper tracings of the waveforms. By the 1990s digital and microprocessor systems allowed preset testing protocols, automatic cursor assignment (used to identify different components of the wave form) and report generation. Currently, nearly all NCS devices have some elements of embedded automation. Additionally, in many instances a technician will perform the test, with cursory physician review of the printed wave forms followed by interpretation. Essentially automation represents a spectrum, with manual performance of all aspects of the test at one end, ranging to an automated test requiring limited human oversight at the other end.[2]

In 1999, the NC-stat device received its original FDA clearance for marketing through the 510(k) process. The most recent FDA clearance was obtained in 2006 (K060584). The predicate device for the FDA review was a traditional NCS device (TECA TD-10/TD-20), and based on the data submitted, the FDA concluded that the technical performance of the two devices was similar. The NC-stat device is designed to facilitate standardized electrode placement and to automate the NCS testing up to physician interpretation/diagnosis. Table 1 summarizes the distinguishing features of the NC-stat device compared to other commonly used NCS devices.

Table 1. Distinguishing Features of NC-stat and Traditional Nerve Conduction Study Devices

	Traditional NCS Equipment	NC-stat Equipment	Comment
Electrode placement	Requires manual placement of electrodes, based on anatomical landmarks.	Uses an integrated electrode configuration of electrodes, tailor made for median, ulnar, peroneal, tibial and sural nerves. Placement of integrated electrode components is based on anatomical landmarks.	NC-stat device designed to ensure reproducible placement of electrodes, suitable for health care personal with limited training. Similar in concept to electrode arrays used for EKGs. Available NC-stat electrode configurations currently precludes assessment of children and also some nerves, such as medial and lateral antibrachial cutaneous nerves of the arms, lateral femoral cutaneous nerve of the thigh and saphenous nerve of the lower leg. Traditional NCS equipment allows for individualized electrode positioning by the examiner and requires specialized training.
Temperature control	Temperature measured by clinical assessment or surface thermometer.	Device includes an embedded digital thermometer allowing normalization of the data to standard temperatures.	Temperature can affect waveforms, and high or low temperatures require correction using linear regression equations. NC-stat device will not record if temperature is outside a given range, and the user will be instructed to warm or cool the limb.
Identification of supramaximal stimulus	Stimulus intensity manually increased while observing the evoked response.	Using device software, the supramaximal stimulus intensity is determined using a sequence of electrical stimuli based on a priori knowledge of nerve tested and patient characteristics.	The supramaximal stimulus intensity is identified by a plateau in the amplitude of the motor or sensory evoked response. The supramaximal stimulus intensity varies according to the nerve tested and patient characteristics, such as BMI. In the NC-stat these parameters are embedded in the software. NC-stat designed to minimize false readings by ignoring waves that do not increase with increasing stimulus intensity.
Acquisition of waveforms	Automatically or manually change recording parameters for acquisition of different waveforms.	Automatically change recording parameters for acquisition of different waveforms	Filter, gain (i.e. adjusting sensitivity), timebase and other parameters vary according to whether recording CMAP, SNAP or F-waves. For example, due to its low signal to noise ratio, the sensory waveform must often be averaged to eliminate artifacts.

Quality checks	Throughout data acquisition, technician or physician must evaluate waveforms to determine whether they are physiologically reliable.	Automated quality control checks include detection of saturation, excessive stimulus artifact, noise levels and unexpected changes in waveform morphology.	If quality checks fail, NC-stat will not report data.
Cursor assignment	Manually placed or automated. Automated measurements are commonly determined by identification of deflections from baseline that exceed a preset threshold and are in the correct direction. .	Automated cursor assignment designed to incorporate the complexities of the stimulus artifact, baseline variation, noise and complex wave morphology.	Cursor assignment identifies the various components of the waveforms and allows calculation of latency, amplitude, conduction velocity, etc. Cursor assignment routinely automated in traditional NCS tests, similar in concept to automated EKGs where different components of P wave, QRS complex and T waves are identified and measured. Traditional automated cursor assignment functions well only if the tracing is flat, and free of contamination and noise.
F wave analysis	Either manual assignment or automated, based on simple algorithm whereby first deflection from the trace baseline exceeding a preset threshold is identified as the latency.	Designed to provide more sophisticated F wave analysis. For example, other late-wave components (A-wave and M-wave) are removed and noise removed by filtering. Windows are defined for F-wave latency search based on the expected range of F wave latencies from reference data.	F waves are low amplitude, inherently variable responses that may not be detected after each stimulus. Therefore, analysis of F-waves requires evaluating a series of responses rather than a single response. F wave latency is the most commonly reported F wave parameter, and is reported as a minimum, median or mean value for a series of F waves. Automated F-wave in traditional NCS is limited if the tracing is not flat or if the F-wave morphology is complicated and poorly differentiated from the trace baseline. Manual correction may be performed with traditional NCS.
Comparison of NCS parameters to normative/reference ranges	Normative/reference range data should be based on individual lab experience, but instead is often adapted from published reference ranges that approximate local settings.	Reference ranges based on data using identical methodology that is embedded in the device. Reference range adjusted according to the patient's gender, age, height and weight.	References ranges may introduce a source of bias, and may be partially responsible for interobserver variability. For NC-stat, a consistent reference range is used across all labs.

As can be seen from Table 1, any NCS is a multistep process involving test set up, data acquisition and data analysis and thus the accuracy of the test depends on adherence to the many technical details. Minor variations in any one of these steps can result in varying results. The additional automation offered by the NC-stat device is designed to minimize the accumulation of minor technical differences related to human interactions, i.e. those differences that contribute to both intra- and interexaminer variability. Although the concept of further automation of NCS studies is inherently appealing, prior to its acceptance, results of the NC-stat must be assessed for test-retest reliability and then compared to traditional methods.

Assessment of diagnostic devices is often focused on three sequential parameters; 1. technical performance; 2 diagnostic performance; and 3. demonstration that the diagnostic information can be used to improve patient outcomes. This approach is appropriate for novel diagnostic technologies. For example, for a novel diagnostic test, such as an imaging test, it is important to not only know whether the technology can image the target anatomy, but how this imaging compares to other unique imaging techniques, and ultimately whether this diagnostic information can be used to improve patient management. The NC-stat device is not a novel technology, but an enhancement of existing NCS technology. The role of nerve conduction studies in the evaluation of patients with symptoms suggestive of neuropathy has already been established, and therefore once the technical performance of the NC-stat has been established, both its diagnostic performance and its role in patient management can be extrapolated from the established role of NCS in general. Similarly, "traditional" NCS technology has undergone a variety of enhancements over the years, representing incremental automation in both the test set up and interpretation; the technical performance associated with each of these enhancements has always been the key parameter. The NC-stat merely represents a highly automated version of a NCS, based on the same physiologic principles as prior generations of NCS devices.

The automated features of the NC-stat permit the performance of NCS outside the neurophysiology laboratory. For example the NC-stat device has been referred to as a "point of care," "handheld" or "portable" device. However, these features are not unique to the NC-stat device and may be a component of what is considered a "traditional" NCS device. Indeed, since all NCS devices exist on a continuum of automation and computerization, it is difficult to precisely differentiate the NC-stat device from all other NCS devices. For example, the term "automated point-of-care device" does not uniquely define the NC-stat device. At best, it may describe the clinical setting within which nerve conduction studies are performed.

While some neurologists have received specific training in performing and interpreting NCS, a wide variety of physicians routinely perform NCS, including physiatrists, orthopedic surgeons, occupational medicine physicians, internists and other primary care physicians. The highly automated features of the NC-stat device may encourage a broader range of physicians to offer NCS. This raises the question of whether physicians without extensive training in neurophysiology can adequately interpret the results of an NCS. However, this is a credentialing issue that is relevant to any type of NCS study, whether it is performed with the NC-stat or other NCS equipment. The increasing automation of EKGs offers an analogy. In their infancy, EKGs were primarily within the purview of the cardiologist, but with increasing automation, both in the placement of electrodes, but also with automated read-outs, EKGs are performed and interpreted by a broad spectrum of physicians.

Therefore, the following technology assessment focuses on the comparative technical performance of the NC-stat and traditional NCS devices, focusing on NCS studies of the extremities. The comparative diagnostic performance of the two different types of devices is not considered relevant to the validation of the NC-stat as an alternative to traditional NCS devices, since both categories of device are based on the same neurophysiologic principles.

Methods

A search of the MEDLINE database was performed for the period of 1986 to July 2008 using the subject heading "NC-stat." This search retrieved 14 articles. In addition, the bibliographies of these studies were reviewed to identify relevant citations not identified in the MEDLINE search. A bibliography supplied by the manufacturer of the NC-stat device (Neurometrix, Waltham MA) was also reviewed. The search identified 21 articles in the published peer-reviewed literature and 17 abstracts and meeting presentations.

Full-length articles which met the following inclusion criteria were included in the initial review:

- Peer-reviewed published full length article published in the English language.
- Prospective enrollment of patients undergoing NCS of the upper or lower extremity.
- Minimum sample size of 10 patients.
- Study focused on either test-retest repeatability of the NC-stat or compared the technical performance of NC-stat and traditional methods in the same patients.
- Analysis of the two electrodiagnostic studies was blinded to one another.
- Statistical analysis included data correlation between the two tests (i.e. Pearson correlation coefficient, Spearman rank order coefficient or intraclass correlation coefficient (ICC).

Using these criteria, 10 studies were identified for initial review. One study, Vinik et al, was eliminated from final review, since interpretation of studies was not blinded.[3] A study by Katz investigated whether the NC-stat device could be used as a screening tool to identify carpal tunnel syndrome in a population of 1695 asymptomatic industrial workers.[4] This study was eliminated from final review because there was no direct comparison to traditional NCS. This study focused on the definition of normal values defined by the NC-stat device compared to the distribution of results in industrial workers. The author questioned the NC-stat reference values, however, use of industrial workers to define normal values is questionable. Of the articles selected for final review, one study focused on test-retest parameters of the NC-stat device, and the remaining 8 studies compared the technical performance of the NC-stat device and traditional methods. Studies included patients with a variety of disorders, including carpal tunnel syndrome or diabetes, two of the most common indications for nerve conduction studies.

Test-Retest of NC-stat

Kong and colleagues studied the test-retest characteristics of the NC-stat device in 21 healthy volunteers who underwent testing of the upper or lower extremity, or both within a 7 day period, performed by the same technician.[5] There were 15 NC-stat studies performed on the upper extremity in one group and repeated within a week. Nine of these subjects had additional studies performed on the lower extremity, and 6 subjects had studies performed only on the lower extremity. The lower extremity studies were also repeated within the week. Distal motor latency (DML), mean F-wave latency (FWL), distal sensory latency (DSL), compound muscle action potential (CMAP) amplitude, and sensory nerve action potential (SNAP) amplitude were recorded from the median, ulnar, peroneal and tibial nerves as appropriate. The following statistical analyses were performed on the test- retest data: Pearson correlation coefficient (CC) to assess the association between recorded parameters obtained 7 days apart; intraclass correlation coefficient (ICC) to determine the agreement between the paired tests and relative intertrial variation (RIV) was used to assess data variability. Of these, the ICC is the most relevant, since this statistic has also been reported in test-retest studies of traditional NCS studies, allowing comparison.

The ICC for all parameters tested ranged between 0.8 and 0.9 with three exceptions. The ICC for median F-wave latency was 0.68, for peroneal CMAP 0.33 and for tibial CMAP the ICC was 0.73. The patterns of reproducibility in this study were similar to traditional methods. For example, as discussed further below, for both NC-stat and traditional methods, F-wave latency has the best reproducibility, followed by DML, DSL and SNAP amplitude. CMAP amplitude had the lowest reproducibility, with peroneal CMAP the lowest. Limitations of this study include the small sample size and the use of healthy volunteers, which does not duplicate the intended population with symptoms suggestive of neuropathy.

Comparative Studies of NC-stat and Traditional Methods

Upper Extremity

Leffler and colleagues reported on the comparative diagnostic performance of NC-stat and traditional methods in 150 consecutive patients referred for evaluation of upper extremity or neck symptoms.[6] Results in the first 75 patients were used to refine the signal processing algorithms of the NC-stat device, specifically focusing on the stimulation patterns and algorithms used to record and calculate the F-wave latencies. The second group of 75 patients was considered the validation group. In the validation group, the Pearson correlation coefficient for distal motor latency (DML) was 0.94, and 0.86 for median F-wave latency.

Median and ulnar distal sensory latencies (DSL) and amplitudes (SNAP) were recorded using traditional methods and NC-stat in 60 consecutive subjects referred for evaluation for upper extremity or neck symptoms.[7] The median-ulnar DSL difference (MUD), a key parameter in the evaluation of carpal tunnel syndrome, was also calculated. The study included patients with a wide spectrum of neuropathology and thus was not limited to those with suspected carpal tunnel syndrome. The statistical analysis included paired t-test, Pearson correlation coefficient and intraclass correlation coefficient. The ICC was 0.69 for the ulnar DSL, 0.91 for the median DSL and 0.87 for the MUD measurement. Differences between the techniques of amplitude measurement (baseline-peak vs. peak-peak) precluded the use of ICC in comparing measurements of amplitude. Therefore, for these outcomes only the Pearson correlation coefficient was reported. The Pearson correlation coefficient was 0.83 and 0.88 for the ulnar and median SNAP amplitudes respectively. Statistically significant systematic differences, as measured by the paired t-test, were observed between the mean values for all parameters tested, with the NC-stat latencies longer than those recorded with traditional equipment. These differences are expected and related to the different placement of electrodes between the two tests. Separate reference values are therefore required for NC-stat and for traditional NCS studies.

Armstrong and colleagues studied 33 patients referred for electrodiagnostic testing of the upper extremity, and thus this population was not limited to those with suspected carpal tunnel syndrome and a larger number of nerve conduction parameters were reported.[8] The Pearson correlation coefficient was 0.91 and 0.79 for median nerve distal motor and sensory latency, respectively, which the authors considered consistent with other studies. Similarly, the correlation coefficients for median to ulnar DML and DSL differences were 0.88 and 0.82, respectively. However the correlations for the ulnar distal motor and sensory latencies was considerably lower at only 0.40. The authors suggest that this outlier value may be related to the fact that all the ulnar latencies fell within a small range due to a low prevalence of ulnar pathology. Therefore, the correlations were lower than they might have been if the study population had a wider range of abnormal values. The correlation between NC-stat and traditional methods is limited by the baseline test-retest variability and interexaminer variability of the traditional control group. The authors of this study point out that the correlation coefficients in this study are similar to the test-retest correlation for traditional methods in a study by Salerno[11] (discussed further below). That study also reported lower correlation coefficients for ulnar nerve measurements compared to median nerves.

Two studies have specifically compared the results of NC-stat and traditional NCS in patients with known or suspected carpal tunnel syndrome. For example, Rotman and colleagues analyzed median nerve distal motor latencies (DML) with traditional methods and NC-stat in a group of 46 patients (88 hands) with known carpal tunnel syndrome scheduled to undergo surgery.[9] The Pearson correlation coefficient between the 2 DML measurements was 0.94. Similarly Elkowitz and colleagues studied 72 patients with carpal tunnel syndrome who had undergone NCS by both traditional methods and NC-stat prior to surgery.[9] These patients were drawn from a larger database of 350 patients with carpal tunnel syndrome. The Pearson correlation coefficient for the median nerve distal motor latency was 0.88.

Lower Extremity

Perkins and colleagues compared the diagnostic performance of traditional methods and NC-stat in the evaluation of sural nerve amplitude potentials in a consecutive case series of 72 patients with diabetes.[12] Correlation was analyzed using Spearman's correlation coefficients for all comparisons. Agreement for sural nerve amplitude potential values between the two techniques was evaluated using Bland Altman analysis. The Spearman correlation coefficient was 0.95. A Bland Altman analysis revealed a small systematic bias, such that the NC-stat device underestimated the sural nerve potential by 1.2 μV compared to traditional methods. This bias is explained by the fact that the NC-stat device reports sensory amplitudes less than 2.1 μV as absent. However, sensory amplitudes below this level are consistent with a neuropathy, similar to an absent signal.

In a follow-up publication, the authors reported on additional assessment of the median and peroneal nerve.[13] The Spearman correlation coefficient was high for all parameters studied, ranging from 0.90 for the sensory amplitude of the median nerve, to 0.76 for the motor F-wave latency. Similar to the analysis of the sural nerve, a Bland Altman analysis revealed a small systematic bias. To determine the impact of this systematic bias, the authors reported that of the 50 patients with diabetic neuropathy, 44 would have been classified appropriately using the NC-stat device, resulting in a sensitivity of 88%. Of the 22 without neuropathy, 18 were classified appropriately, resulting in a specificity of 82%. The authors further pointed out that the 10 misclassifications had minor electrophyioslogical abnormalities. These estimations are limited by the assumption that results of traditional NCS are accurate, an assumption that is challenged by the results of test-retest reliability, discussed further below.

Peroneal and posterior tibial distal motor latency (DML), amplitude (CMAP) and F-wave latency (FWL) were obtained with both traditional NCS methods and the NC-stat device in a series of 60 patients referred for a lower extremity NCS.[14] Both studies were performed by the same technician. Correlation between the two measures was assessed with the Spearman correlation coefficient. The Spearman correlation was considered the key outcome since it provides a rank order assessment of paired variables and thus controls for systematic biases due to methodologic difference between the two techniques. Based on prior studies of traditional methods, acceptable correlation was defined as 0.60 to 0.79, and an excellent correlation was defined as ≥ 0.80. Correlation was considered excellent for peroneal and posterior tibial FWLs and the peroneal CMAP. Acceptable correlation was identified for peroneal DML and posterior tibial CMAP. The Spearman correlation coefficient for posterior tibial DML was only 0.45.

Systematic bias between the two measurements was assessed by comparing mean measurements using a paired t-test. There was no statistically significant difference between the two techniques for mean peroneal CMAP and posterior tibial FW latency. All the other parameters had differences in their mean measurements. This finding does not suggest that the results of the NC-stat test are invalid, rather that different reference ranges are needed for NC-stat compared to the methods employed by this reference laboratory. For example, normal reference ranges are theoretically established by individual laboratories performing traditional NCS. In contrast, the increased automation associated with the NC-stat device permits the use of a consistent reference range across all laboratories.

The authors also performed a sub-group analysis comprised of the 50% most severely affected nerves. The number of patients and nerves in the analysis was not provided. An acceptable Spearman correlation coefficient was maintained for all parameters except for the DML in the both the peroneal and tibial nerves. This finding is difficult to interpret since severely diseased nerves show more variable responses which lead to increased subjectivity in interpreting the results.

Table 1. Summary of Correlation Coefficients* of Comparative Studies

Study	No. Pts	DML	CMAP Amplitude	F-wave Latency	DSL	SNAP Amplitude
§Leffler[6]	75	(m) 0.94		(m) 0.86		
§Rotman[8]	46	(m) 0.94				
§Kong[7]	60				(m) 0.91 (u) 0.70 (mu) 0.88	(m) 0.88 (u) 0.83
Elkowitz[9]	72	(m) 0.88				
¶Jabre[14]	60	(p) 0.70† (t). 0.45†	(p) 0.86† (t) 0.73†	(p) 0.91† (t) 0.90†		
¶Perkins[12]	72					(s) 0.95†
¶Perkins[13]	72	(m) 0.83† (p) 0.83†		(m) 0.76† (p) 0.86†	(m) 0.83† (p) 0.86†	
Armstrong[10]	33	(m) 0.91 (u) 0.40 (mu) 0.88			(m) 0.79 (u) 0.40 (mu) 0.82	

*Pearson correlation coefficient unless otherwise specified
†Spearman correlation coefficient
m, median nerve; u, ulnar nerve; mu, median-ulnar difference; p, deep peroneal; t, posterior tibial; s, sural.
§ Publication includes NeuroMetrix authors
¶NeuroMetrix provided funding

Discussion

All reviewed studies report high correlation for the parameters studied with few exceptions. For example, in the test-retest study of Kong, the ICC for all parameters tested ranged between 0.8 and 0.9 with the exception of median F-wave latency (0.68), peroneal CMAP (0.33) and tibial CMAP (0.73).[5] As noted in the above table, in studies comparing NC-stat and traditional methods, the correlation coefficient (either Pearson or Spearman) ranged between 0.70 and 0.95, with the exception of 0.45 for DML of the tibial nerve. The strong correlation between parameters is not surprising, given that the physiologic basis of the two studies is identical. Essentially, the NC-stat device is intended to duplicate traditional methods, with the added advantage of additional elements of automation, designed to eliminate the numerous small technical errors that cumulatively can contribute to study variability.

In this analysis, traditional NCS methods are considered the gold standard, and thus it is instructive to review the reproducibility of the gold standard for comparison purposes. Due to the many technical variables inherent in traditional NCS studies, it is expected that test-retest studies of traditional NCS will show some variability, which will then impact on the correlation with NC-stat. In addition all NCS studies are data-rich with multiple different parameters recorded for multiple different nerves, representing different technical challenges. Therefore, the reproducibility of any NCS will vary with the parameter studied. In general, the reproducibility of NCS of the upper extremity is better than the lower extremity. For example, the tibial nerve innervates multiple small muscles of the foot. Since tibial latencies are based on waveforms recorded between the ankle and foot, any slight changes in placement of the electrodes can record from different muscles, resulting in variable results of nerve latencies. Waveforms from the sural nerve tend to be very small and difficult to record, and the signal can be further deteriorated by the presence of edema. In general, reproducibility of F waves is higher than other parameters. For example, F waves of the tibial nerve are based on nerve signals between the ankle and spine, thus minimizing the contribution of the variable component between the ankle and foot.

Several large test-retest studies of traditional methods illustrate the varying reproducibility of the different components of NCS. Bril and colleagues investigated the test-retest reliability of NCS in a diabetic neuropathy trial using an explicitly designed NCS protocol that that required interpretation of all studies by a central core laboratory.[15] The purpose of the investigation was to maximize the reproducibility of nerve conduction studies as used as a research tool; highly reproducible NCS enhance their sensitivity in detecting a treatment response and thus will increase the power of the studies that use NCS as an efficacy measure. A total 1345 patients with mild diabetic neuropathy at 60 international centers underwent triplicate NCS studies at the study onset. Results were presented as the coefficient of variation (COV), where an increasing value reflects increasing variability. The COVs for different parameters ranged from 3-17%. The highest COVs were seen in the measurement of amplitudes of different nerves. In contrast conduction velocities and distal latencies showed lower levels of variability. Bird and colleagues performed a similar study of over 1100 participants with mild diabetic neuropathy participating in a clinical trial of the investigational drug zenarestat.[18] NCS studies were repeated in triplicate on separate days within a 4 week window by a certified electromyographer and reviewed at a central laboratory.[16] Results were reported in terms of the coefficient of variation. The patterns of the COV were similar between the two studies, with the highest values reported for amplitude measures and lowest for nerve conduction velocity measures. However the level of variation was higher in this study compared to earlier Bril study,[15] ranging from 52-53% for amplitude testing to a low of 7% for nerve conduction velocity of the median nerve.

Another measure of technical performance is the percentage of NCS that meet testing guidelines. This parameter was not reported in the comparative studies of the NC-stat device, but was reported in a retrospective analysis of registry data of the NC-stat device used for the evaluation of carpal tunnel syndrome.[17] The registry data set included the results of NCS using NC-stat in 1190 patients performed by 613 different physicians representing a range of medical specialties. Two levels of compliance to testing guidelines were defined. A NCS was considered strictly compliant if at least one limb with both median and ulnar distal sensory measurements were reported. A less restrictive compliance level was defined as two or more upper extremity distal sensory measurements. A total of 81.6% of studies met the strict definition and 92.6% met the less restrictive definition.

These registry results can be compared to a study of testing compliance to traditional methods reported by Storm and colleagues. This study based on a Medicare insurance claims database of 1567 patients with carpal tunnel syndrome who went on to carpal tunnel release.[18] The authors defined an inadequate study as less than 2 sensory nerve conduction studies, which is consistent with practice standards published by the American Academy of Electrodiagnostic Medicine.[19] A total of 90.1% had at least two sensory nerves studied, comparable to the 92.6% compliance reported in the NC-stat database. Results of other studies analyzing compliance to testing standards also report that 100% compliance is not readily achievable for this technically demanding test. For example in the study by Bril and colleagues reviewed above, the core laboratory had comments, corrections or rejections on 34.3% of submitted traditional NCS studies.[15] Similarly, in the Bird study,[16] corrections or requests for repeat studies were reported in 19% of cases.

These studies illustrate the challenge in standardizing traditional NCS, and by extension, the challenge in comparing results of the NC-stat device and traditional methods. The inherent challenge is further complicated by the variety of statistical tests that may be reported. There is no standard statistical method for comparing the results of electrodiagnostic tests, and studies have reported outcome in terms of intraclass correlation coefficients (ICC), Pearson and Spearman correlation coefficients. ICC is typically calculated when no bias between the two methods is anticipated, and thus ICC can be used in test-retest studies of NCS as performed with the same equipment by the same operator. The Pearson correlation coefficient controls for any bias in two measurements, and thus is commonly used when comparing two related measures. The Spearman correlation coefficient is a variant of the Pearson correlation coefficient and evaluates results in terms of their rank order, as opposed to the absolute value. If there is no source of bias, the ICC should be very similar to the Pearson or Spearman correlation. If there is a source of bias, then the ICC will be lower than other correlation coefficients. As noted in the introductory section, there are multiple sources of potential bias between laboratories performing NCS and results of NCS studies must be compared with reference ranges that are derived from individual laboratories. One common source of bias is placement of electrodes, which may vary among laboratories. In contrast, the NC-stat device is designed to reduce sources of bias. For example, placement of electrodes is highly standardized in the NC-stat device, and reference values for NC-stat results are not laboratory dependent, but are derived from a large database.

It must be recognized that due to its inherent variability, NCS studies as performed by multiple laboratories are a flawed gold standard, such that the reported correlation coefficients of NC-stat vs. traditional methods reflect the limitations of the traditional methods as well as the comparative performance of an NC-stat study. Essentially, correlation studies between NC-stat and traditional NCS can only be reasonably expected to demonstrate up to the level of reliability shown in test-retest studies of traditional methods. Correlation coefficients from test-retest studies using traditional methods can be compared to the comparative studies of NC-stat and traditional methods to provide an estimation of the diagnostic performance of NC-stat compared to the gold standard. Table 2 summarizes the correlation coefficients for common NCS parameters reported in comparative and test-retest studies. While different types of correlation coefficients (e.g. ICC, Pearson, Spearman) are used among studies, it is reasonable to assume that the ICC reported for test-retest studies should be very similar to the Pearson or Spearman coefficients, since there should be no source of bias in test-retest studies using the same equipment and technique.

Table 2 also provides a comparison of correlation coefficients of test-retest studies of traditional methods and comparative studies of NC-stat and traditional methods. The study by Kohara and colleagues reported the reproducibility of traditional methods in 132 healthy subjects and 172 patients with diabetic neuropathy who underwent 2 NCS studies within a 1 to 4 week period by the same examiner.[20] Results were presented separately for healthy patients and those with diabetic neuropathy. Table 2 only includes the results for the 172 patients with diabetic neuropathy. Salerno and colleagues studied the inter- and intra- examiner reliability of traditional methods in 161 keyboard operators.[11] Participants underwent sensory NCS of the both wrist by two different examiners at one setting, and then repeat NCS 3 weeks later by one examiner. The Salerno study reports similar results for both inter and intra-examiner and for the dominant and nondominant hand. Only the representative results for the interexaminer dominant hand are presented in Table 2. The study by Bird, reviewed above, reported results both in terms of coefficient of variation and ICC.[16] The ICC results are presented in Table 2 and are reported as the range of ICC values at baseline for one to three replicates for all treatment groups. The study by Dyck and colleagues the reproducibility of various nerve conduction parameters was assessed in 20 subjects with varying degrees of neuropathy.[21] The subjects were tested 3-5 days apart and reproducibility was reported using the ICC. Finally the most recent study by Pinhiero and colleagues reported the ICC values in 32 healthy volunteers who underwent repeat testing at an interval of 27-872 days.[22]

Table 2. Summary of Comparative Studies of Nerve Conduction Technologies

	NC-stat vs. Traditional Methods	Test-Retest of NC-stat	Test-Retest of Traditional Methods				
Parameter	Results summarized from Table 1 (n=418)	Kong[5] (n=21)	Kohara[20] (n=172)	Salerno[11] (n=161)	Bird[16] (n=1100)	§Dyck[21] (n=20)	Pinheiro[22] (n=32)
<u>DML</u>							
median	0.83-0.94	0.87^	0.95^			0.82^	
ulnar	0.40	0.83^				0.70^	0.34^
peroneal	0.70	0.85^				0.75^	
tibial	0.45	0.82^	0.66^			0.62^	0.36^
<u>CMAP</u>							
median		0.86^				0.65^	
ulnar		0.97^					0.64^
peroneal	0.86	0.33^				0.98^	
tibial	0.73	0.73^	0.89^			0.95^	0.68^
<u>F-wave</u>							
median	0.86	0.68^	0.93^		0.87-0.90^	0.52^	
ulnar		0.92^				0.85^	0.89^
peroneal	0.86-0.91	0.90^			0.83-0.89^		
tibial	0.91	0.94^	0.93^			0.91^	0.95^
<u>DSL</u>							
median	0.79-0.91	0.88^		0.91		0.89^	0.81^
ulnar	0.70	0.80^		0.64		0.76^	0.50^
median-ulnar	0.88			0.89			
sural						0.75^	0.63^
<u>SNAP</u>							
median	0.88	0.95^	0.91^	0.87	0.91-0.92	0.93^	0.77^
ulnar	0.83	0.98^		0.85		0.93^	0.55^
sural	0.95				0.76-0.80	0.93^	0.74^

*Unless otherwise indicated Pearson or Spearman correlation coefficient is indicated; ^intraclass correlation coefficient (ICC).
†median, median nerve; ulnar, ulnar nerve; peroneal, deep peroneal nerve; tibial, posterior tibial nerve; sural, sural nerve; DML, distal motor latency; CMAP, compound muscle action potential amplitude; F-wave, F-wave latency; DSL, distal sensory latency; SNAP, sensory nerve action potential amplitude.
§Correlation coefficients estimated from Figure 6.

As noted in the above table, results of the NC-stat test-retest, and studies comparing NC-stat with traditional methods, show similar patterns of correlation coefficients for different NCS parameters, and generally in the same range of 0.80-0.90. For example, the correlation coefficient for tibial DML are the lowest for comparative NC-stat study (0.45) and test-retest of traditional methods (0.62-0.66), consistent with technical challenge of measuring this parameter. In contrast the correlation coefficients are the highest for measures of F-wave latencies. The correlation coefficients for peroneal CMAP (0.33) and median F waves (0.68) reported by Kong in the test-retest study of NC-stat appear to be outliers to the overall pattern of results.[5] The explanation of this finding is not readily apparent, but may be related to the small number of patients (n=20) in the study. The generally lower ICC values reported in the Pinhiero study are also not readily explained, although the variability in results does highlight the challenges and limitations in comparing the results of the NC-stat test to a variable standard.

Needle electromyography (EMG) cannot be performed with the current version of the NC-stat device. A position statement by the American Association of Neuromuscular and Electrodiagnostic Medicine (AANEM) states that in most cases an electrodiagnostic consultation involves using both NCS and needle EMG.[23] This position statement is largely unreferenced and does not describe the process underlying the recommendations. Therefore, the evidence basis for the recommendation of combined NCS and EMG studies is not apparent. However, needle EMG is not needed in some research and clinical situations. For example, a focused NCS study has been used to evaluate diabetic neuropathy. The NC-stat device has been incorporated into clinical trials of drug therapy for diabetic neuropathy. In one study of duloxetine enrolling 1,000 patients, results of the NC-stat were incorporated into the FDA approval process. NC-stat was used to evaluate the ulnar and peroneal nerves to determine whether the drug degraded motor and large sensory fiber function.[24] Using the NC-stat device to perform peroneal nerve conduction studies, the rate of diabetic neuropathy and association to clinical factors was equivalent to other research studies of diabetic neuropathy using traditional NCS. Vinik and colleagues have also reported that the NC-stat can be used at the point of service to evaluate diabetic neuropathy.[25]

Evaluation of carpal tunnel syndrome typically does not require an EMG; physicians, including neurologists, physiatrists or hand surgeons, may focus on demonstrating a focal median neuropathy at the wrist and do not perform an EMG. A group of specialty societies (including the AANEM) published a practice parameter addressing electrodiagnostic tests for patients with suspected carpal tunnel syndrome.[18] This document identified various parameters of an NCS study as “standard,” but indicated that an EMG is only considered an “option.” A “standard” is defined as “generally accepted principles for patient management” while an “option” is defined as “other strategies for patient management for which the clinical utility is uncertain.” The American College of Occupational and Environmental Medicine has issued guidelines regarding elbow complaints and a separate guideline on forearm wrist and hand complaints. For elbow complaints, NCS studies are recommended to confirm ulnar nerve entrapment if conservative measures fail, while EMG studies are considered an option in the small subset of patients to distinguish lateral epicondylitis from radial neuropathy.[26] For forearm, wrist and hand complaints, NCS are recommended to assess median and ulnar nerve impingement after failure of conservative therapy.[27] Needle EMG studies are not suggested (either as a recommendation or as an option). Notably these guidelines indicate that the conclusions are based on category "D" evidence, defined as "panel interpretation of information not meeting inclusion criteria for research based-evidence or consensus." The practice parameters and guidelines reviewed above are all based on consensus opinion and it is apparent that there is inadequate evidence from controlled trials to specifically establish which components of an NCS or EMG should invariably be part of an electrodiagnostic study for a given diagnosis. In part the lack of quality evidence is a result of the many different possible combinations of parameters of both and NCS and EMG, suggesting that the physician must tailor the electrodiagnostic exam to the specific clinical situation.

Summary

In summary, the data from 8 studies comparing the technical performance of NC-stat and traditional methods showed excellent correlation that is similar in pattern and range for test-retest studies using traditional methods. Three studies included NeuroMetrix (manufacturer of the NC-stat device) authors, three studies included some level of funding by NeuroMetrix as is typical in the medical device industry, and two studies lacked corporate sponsorship. These results support the use of the NC-stat device as an alternative to traditional methods for NCS studies. The strong correlation between these two categories of devices is not surprising since both are based on identical neurophysiologic principles, and the NC-stat device is essentially a highly automated version of the traditional NCS device.

Glossary of Terms

Amplitude: The height of an action potential measured from baseline to peak or peak to peak. The motor amplitude (i.e. CMAP, see below) is measured in millivolts (mV), and the sensory nerve amplitude (SNAP, see below) is measured in microvolts (μV).

Coefficient of variation (COV): A statistical representation of the precision of a test, defined as the ratio of standard deviation to its arithmetic mean, typically reported as a percentage. As the COV increases, the repeat measurements have greater variation around their mean. As a hypothetical example, a COV of 5% would imply that approximately 95% of repeated measurements would be within 10% of the mean value.

Compound muscle action potential (CMAP): CMAP represents the sum of the surface recorded motor unit action potentials. Measured parameters include distal latency, amplitude and conduction velocity. The initial response of a muscle to an electrical stimulus, resulting from the action potential traveling from the point of stimulation to the muscle. This initial response is the largest in amplitude.

Conduction velocity: Speed of electrical conduction along a nerve measured in meters/second.

Cursor assignment: Identification of key points on the waveform, such as onset and peak.

F-waves: Electrical stimulation of a motor nerve results in action potentials that travel both proximally and distally. For example, the CMAP results from the stimulation proceeding directly from the nerve to the muscle. The smaller F waves result from the action potential which travels first proximally to the motor neuron cell bodies in the spinal cord, and then traveling back distally down the nerve fibers to stimulate the muscles after a brief delay. F waves are referred to as a late response (see below).

H waves: H waves are the electrical equivalent of deep tendon reflexes, and are also considered a late response (see below). In clinical practice, H-reflex testing of the soleus, gastrocnemius and flexor carpi radialis may be used.

Intraclass correlation coefficient (ICC): The ICC is based on analysis of variance techniques and is used to assess the agreement between two tests. The ICC takes a value of 0 to 1 and is defined as the ratio of the within-measurement variance to the total variance derived from the within- and between- measurement variance. If the between-measurement variance is low (i.e. high agreement) then the ICC will approach 1. If the between-measurement variance is high relative to within-measurement variance then the ICC will approach 0. The ICC is sensitive to systematic differences between two methods and is therefore preferred when comparing methods expected to generate numerically identical outputs.

Late responses: Collectively refers to the secondary responses that are seen after the initial, large amplitude response, and refers to both F waves and H waves. This type of late response is useful to evaluate the integrity of more proximal nerves that are inaccessible to direct assessment, and can assist in the evaluation of radiculopathies, plexopathies and polyneuropathies. Assessment of late responses requires different machine settings and set up compared to CMAP.

Latency: A measurement of the time from the moment of stimulation to the appearance of an action potential. Both distal and proximal latency can be measured, depending on the site of the stimulating electrodes. Latency for motor nerves is measured at the onset of the action potential curve, while latency for sensory nerves is measured at the onset or peak of the action potential. Latency is measured in milliseconds.

Paired t tests: Tests whether mean values are the same, and assesses the systematic bias between two measurements.

Pearson correlation coefficient: The most commonly used correlation coefficient for two continuous variables is called the Pearson correlation coefficient. The correlation coefficient is a point estimate for the strength of the association between two continuous variables. The correlation coefficient has a range of values from -1 to +1. The limitation of this statistic is that it measures the strength of a relationship between two variables, and not the numerical agreement between two variables. The Pearson correlation coefficient is not sensitive to bias or differences in scale between the two methods compared.

Sensory Nerve Action Potential (SNAP): The initial response of a sensory nerve to a stimulus. SNAP can be performed either ortho- or antidromically. SNAPs are used to evaluate the function of large myelinated nerves. The two primary measures for SNAP are latency and amplitude. (see sural nerve amplitude, below)

Spearman correlation: A non-parametric measure of correlation that does not assume a linear relationship between variables.

References

1. Barboi AC, Barkhaus PE. Electrodiagnostic testing in neuromuscular disorders. Neurol Clinics 2004;22:619-41.

2. Gozani SN, Fisher MA, Kong X et al. Electrodiagnostic automation: Principle and practice. Phys Med Rehabil Clin N Amer 2005;16:1015-32.

3. Vinik AI, Emley MS, Megerian JT, Gozani SN. Median and ulnar nerve conduction measurements in patients with symptoms of diabetic peripheral neuropathy using the NC-stat system. Diabetes Tech Therap 2004;6:816-24.

4. Katz RT. NC-stat as a screening tool for carpal tunnel syndrome in industrial workers. J Occup Environ Med 2006;48:414-18.

5. Kong X, Lesser EA, Megerian JT, Gozani SN. Repeatability of nerve conduction measurements using automation. J Clin Monit Comput 2006 20:405-10.

6. Leffler CT, Gozani SN, Cros D. Median neuropathy at the wrist: Diagnostic utility of clinical findings and an automated electrodiagnostic device. J Occupational and Environmental Med 2000;42:398-409.

7. Kong X, Gozani SN, Hayes MT, Weinberg DH. NC-stat sensory nerve conduction studies in the median and ulnar nerves of symptomatic patients. Clin Neurophysiol 2006;117:405-13.

8. Rotman MB, Enkvetchakul BV, Megerian JT, Gozani SN. Time course and predictors of median nerve conduction after carpal tunnel release. J Hand Surg 2004;29A:367-72.

9. Elkowitz J, Dubin NH, Richards BE, Shaw Wilgis EF. Clinical utility of portable versus traditional electrodiagnostic testing for diagnosing, evaluating, and treating carpal tunnel syndrome. Amer J Orthop 2005;August:3632-64.

10. Armstrong TN, Dale AM, Al-Lozi et al. Median and ulnar nerve conduction studies at the wrist: Criterion validity of the NC-stat automated device. J Occupational Environmental Med 2008;50:758-64.

11. Salerno DF, Werner RA, Albers JW et al. Reliability of nerve conduction studies among active workers. Muscle Nerve 1999;22:1372-79.

12. 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. Diabetes Care 2006;29:2023-27.

13. Perkins BA, Orszag A, Grewal J et al. Multi-site testing with a point-of-care nerve conduction device can be used in an algorithm to diagnose diabetic sensorimotor polyneuropathy. Diabetes Care 2008;31:522-24.

14. Jabre JF, Salzsieder BT, Gnemi KE. Criterion validity of the NC-stat automated nerve conduction measurement instrument. Physiol Meas 2007;28:95-104.

15. Bril V, Ellison R, Ngo M et al. Electrophysiological monitoring in clinical trials. Roche Neuropathy Study Group. Muscle Nerve 1998;21:1365-7.

16. Bird SJ, Brown MJ, Spino C et al. Value of repeated measures of nerve conduction and quantitative sensory testing in a diabetic neuropathy trial. Muscle Nerve 2006;214-24.

17. Megerian JT, Kong X, Gozani SN. Utility of nerve conduction studies for carpal tunnel syndrome by family medicine, primary care and internal medicine physicians. J Am Board Fam Med 2007;20:60-64.

18. Storm S, Beaver SK, Giardino N et al. Compliance with electrodiagnostic guidelines for patients undergoing carpal tunnel release. Arch Phys Med Rehabil 2005;86:8-11.

19. Jablecki CK, Andary MT, Floeter MK et al. Practice parameter: Electrodiagnostic studies in carpal tunnel syndrome: Report of the American Association of Electrodiagnostic Medicine, American Academy of Neurology, and the American Academy of Physical Medicine and Rehabilitation. Neurology 2002;58:1589-92.

20. Kohara N, Kimura J, Goto Y et al. F wave latency serves as the most reproducible measure in nerve conduction studies of diabetic polyneuoropathy: Multicentre analysis in health subjects and patients with diabetic polyneuropathy. Diabetologia 2000;43:915-21.

21. Dyck PJ, Kratz KM, Lehman KA et al. The Rochester Diabetic Neuropathy Study: design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests. Neurology. 1991;41:799-807.

22. Pinhiero DS, Manzano M, Nobrega JAM. Reproducibility in nerve conduction studies and F-wave analysis. Clinical Neurophysiol 2008. In press.

23. American Association of Electrodiagnostic Medicine. AANEM Position Statement. Proper performance and interpretation of electrodiagnostic studies. Muscle Nerve 2006;33:436-43.

24. Hardy T, Sachson R, Shen S, Armbruster M, Boulton AJ. Does treatment with duloxetine for neuropathic pain impact glycemic control? Diabetes Care. 2007;30:21-6.

25. Vinik AI, Kong X, Megerian JT, Gozani SN. Diabetic nerve conduction abnormalities in the primary care setting. Diabetes Technol Ther. 2006;8:654-62.

26. American College of Occupational and Environmental Medicine (ACOEM). Elbow disorders. Elk Grove Village (IL): American College of Occupational and Environmental Medicine (ACOEM); 2007.
http://www.guideline.gov/summary/summary.aspx?doc_id=10883

27. Forearm, wrist, and hand complaints. Elk Grove Village (IL): American College of Occupational and Environmental Medicine (ACOEM); 2004.
http://www.guideline.gov/summary/summary.aspx?ss=15&doc_id=8545&nbr=4754

Elizabeth Brown, MD. Biography

Dr. Brown is a nationally recognized expert in the field of medical policy development and technology assessment. Dr. Brown has nearly 20 years of experience in evidence based medicine, technology assessment, analysis of medical issues, medical writing, editing and working with panels of physicians. Dr. Brown is presently a Medical Director for Argenta TEC, a health care consulting firm. She was previously Medical Director or Director at Wellpoint Health Networks, Blue Cross Blue Shield Association, Aetna Health Plans, and the American Medical Association. Dr. Brown has published extensively in the area of technology assessment. Dr. Brown is a board certified pathologist.