

JENNIFER NANCY DURRANT

THE RELIABILITY AND VALIDITY OF THE TEN TEST
AND
EXPLORING A NEW VISUAL VERSION

THE RELIABILITY AND VALIDITY OF THE TEN TEST
AND
EXPLORING A NEW VISUAL VERSION

By NANCY DURRANT, B.Sc. PT.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the
Requirements for the
Degree Master of Science

McMaster University © Copyright by Nancy Durrant, March 2014

Descriptive Note

McMaster University MASTER OF REHABILITATION SCIENCE (2014)
Hamilton, Ontario (Rehabilitation Science)

TITLE: The Reliability and Validity of the Ten Test and Exploring a New
Visual Version

AUTHOR: J. Nancy Durrant, B.Sc.P.T. (University of Western Ontario)

SUPERVISOR: Professor J.C. MacDermid

NUMBER OF PAGES: ix, 77

Abstract

Sensation threshold assessment is an important component of physical assessment. Current literature has either limited information on the clinical measurement properties of sensory threshold tests, or has demonstrated concerns in reliability, validity, responsiveness and/or clinical utility. The Ten Test (TT) is an easy and quantifiable test of moving light touch sensation requiring no equipment, however; evidence regarding its reliability and validity are limited. In this thesis, I explored the test-retest reliability and concurrent validity of the Ten Test. I also developed a new, visual version of the Ten Test which was assessed for concurrent validity and patient preferences. The results showed that the Ten Test has excellent test-retest reliability (ICC: 0.83 – 0.91), with acceptable minimal detectable change scores ($MDC_{90} = 1.57 - 2.15$). Ten Test scores did not correlate with current perception threshold or vibration perception threshold scores. The visual version of the Ten Test demonstrated high concurrent validity to the original version of the Ten Test (Spearman's rank correlation coefficient $r = 0.74 - 0.90$), and was preferred by participants (85.7%).

Acknowledgements

It was with the support of some great people that I was able to complete this Master's thesis.

Foremost, I would like to thank my family. Firstly my mother, Susan Durrant, for supporting me throughout my life, for being an example of strength, and for her constant and loving support of myself and my family. My father taught me not to be afraid of big goals or climbing trees. To my husband, Tim Anderson, who always expressed interest in my courses and research, and happily downloaded articles as a recreational pursuit. Also, to my son, Gillis Anderson Durrant, who can't read this now, but in the future can read that he was, and is, a great source of happiness for me.

I would like to express my sincere gratitude to my supervisor, Prof. Joy MacDermid, for her continuous support of my Master's research project, for her patience, motivation, and immense knowledge. Her guidance assisted me through all the time of planning the research and in the writing of this thesis. I would like to express my gratitude to the rest of my thesis committee: Dr. Victoria Galea, and Prof. Anita Gross, for their encouragement, insightful comments, and expert knowledge.

I also thank Margaret Lomotan, who expertly manages the MacHand lab and was always available to support my projects. My fellow graduate students at McMaster University in Rehabilitation Science gave critical insight and suggestions and provided encouragement. A special thank you to my fellow students who made the courses fun and who will always be friends: Dinah Hampson, Diana Last, Jacquie Van Ierssel, and Annie Newman.

Table of Contents

Descriptive Note	iii
Abstract	iv
Acknowledgements	v
List of Figures and Tables	vii
List of Abbreviations and Symbols	viii
DECLARATION OF ACADEMIC ACHIEVEMENT	ix
Chapter 1	1
Introduction	1
A. Description of neuromusculoskeletal conditions	1
B. Description of sensory testing	2
C. Role of sensory evaluation in neuromusculoskeletal assessment	3
D. Clinical Measurement	5
E. Clinical Measurement properties of quantitative sensory tests	8
F. Challenges in the use of sensation assessments	35
Purpose	37
Chapter 2	44
Test-retest reliability and concurrent validity of the Ten Test	44
The validity and patient preference of a visual version of the Ten Test	63
Chapter Three: Future Directions	76

List of Figures and Tables

Tables

Chapter One: Semmes-Weinstein monofilament scale of interpretation	11
Chapter One: Summary of clinical measurement properties of sensory assessments	28
Study One: Table 1: Participant Characteristics	47
Study One: Table 2: Mean scores with standard deviations	51
Study One: Table 3: Test-retest reliability	52
Study One: Table 4: Correlations between the TT, current perception threshold and vibration perception threshold	54
Study Two: Table 1: Participant Characteristics	69
Study Two: Table 2: Descriptive Statistics	69
Study Two: Table 3. Correlations between the TT, current perception threshold and vibration perception threshold	70
Study Two: Table 4. Participant preferences of TT format.....	70

Figures

Study One: Figures 1- 3 Bland and Altman plots with 95% confidence bands.....	53-
.....	54
Study One: Figure 4. The Ten Test.....	56

List of Abbreviations and Symbols

Abbreviations

CI	confidence interval
CPT	current perception threshold
CT	computed tomography
D1	thumb
D2	index finger
D5	little finger
ES	effect size
Hz	hertz
ICC	interclass correlation coefficient
ICC _{2,1}	interclass correlation coefficient: model 2,1
IMHA	Institute of Musculoskeletal Health and Arthritis
LR	likelihood ratio
MDC	minimal detectable change
MDC ₉₀	minimal detectable change at the 90% confidence level
MRI	magnetic resonance imaging
NCV	nerve conduction velocity
PSSD	pressure-specified sensory device
QST	quantitative sensory testing
SD	standard deviation
SEM	standard error of the mean
SWMT	Semmes-Weinstein monofilament test
TT	Ten Test
USD	United States dollars
VisTT	Visual version of the Ten Test
VPT	vibration perception threshold
WEST	Weinstein Enhanced Sensory Test

Symbols

κ	kappa
°C	degrees Celsius
95%CI	95% confidence interval
g/mm ²	grams per millimetre squared
n	number of subjects
<i>p</i>	p-value
<i>r</i>	Pearson's correlation coefficient
<i>r_s</i>	Spearman's rank correlation coefficient

MSc. Thesis – J. Nancy Durrant; McMaster University – Rehabilitation Science

DECLARATION OF ACADEMIC ACHIEVEMENT

The following is a declaration that the content of the research in this document has been completed by Nancy Durrant and recognizes the contributions of Dr. Joy MacDermid, Dr. Victoria Galea and Prof. Anita Gross in both the research process and the completion of the thesis. Nancy Durrant contributed to the study design, was responsible for data collection, data analysis and writing of the manuscripts. Dr. Joy MacDermid conceptualized the study design, and assisted with data analysis and manuscript review. Dr. Victoria Galea and Prof. Anita Gross provided insightful manuscript review.

Chapter 1

Introduction

Chapter one is an overview of current knowledge of sensation threshold testing, as well as a review of reliability and validity. This chapter provides a background literature review on sensory threshold assessment in neuromusculoskeletal disorders, as it applies to the hand. The objective is to determine the clinical measurement properties of the Ten Test have been found, relative to other sensory assessments or outcome measures.

A. Description of neuromusculoskeletal conditions

a. Background. The neuromusculoskeletal system incorporates the muscles, nerves and skeleton of the body. Neuromusculoskeletal disorders are a major burden on Canadians and the health care system. The Institute of Musculoskeletal Health and Arthritis (IMHA) reported that 11 million Canadians were affected by musculoskeletal conditions in 2010, about 3.1% of the population.¹ With the aging population, the IMHA anticipates an increase in the prevalence of musculoskeletal conditions to 15 million Canadians in 2031.¹

The costs associated with musculoskeletal conditions are due to direct treatment costs such as hospital, medication and physician costs, as well as indirect costs such as mortality, and long- and short-term disability.² In the year 2000, the cost of musculoskeletal conditions in Canada was \$22.2 billion, representing the highest total

cost of the top 20 most costly diagnostic categories, followed by cardiovascular diseases and neuropsychiatric conditions.³

b. Role of the nervous system. Although neuromusculoskeletal disorders are often thought of as conditions of bone, muscle and joint, the nervous system plays a critical role in these conditions. Physical function depends on normal integration between sensory and motor systems.⁴ Nerve impairment may impact the functioning of the sensory or motor system, which in turn may impair function.⁵

B. Description of sensory testing

Sensation is the ability to detect and interpret a tactile stimulus.⁶ The process begins with a stimulus to the peripheral sensory receptors.⁷ This triggers action potentials which are transmitted afferently via the peripheral nerves to the dorsal root ganglia.⁷ This is then followed by the dorsal column of the spinal cord to the contralateral cerebral cortex, specifically the somatosensory cortex.⁷ Nerve injury results in degenerative changes throughout the nervous system: to site of the injury, the distal nerve segment, sensory receptors, proximal nerve segment and the cerebral cortex.⁸ For example, in the absence of neural input, sensory end-organs degenerate⁹ and the cortical area mapped by the hand is altered and reduced in size or lost altogether.^{7,8}

Sensory evaluation aids a clinician in several aspects of patient care, such as diagnosing a disorder, identifying the severity of sensory impairment and determining the level of axonal regeneration.¹⁰ Sensory evaluation also helps in determining the best

course of treatment, determining the need for surgical intervention, and in identifying the progression of a condition and establishing the level of hand function.¹⁰ Several sensory tests should be used during an assessment to assess sensibility due to the multiple components contributing to sensory function.

C. Role of sensory evaluation in neuromusculoskeletal assessment

Sensory evaluation has three primary roles: diagnosis, prognosis and outcome evaluation.

a. Diagnosis. Several sensory tests are used for the diagnosis of clinical syndromes. Although nerve conduction velocity (NCV) testing is often employed as a standard test for nerve function, quantitative sensory testing (QST) can detect neuropathy pre-clinically and earlier than NCV tests.¹¹ In conditions such as carpal tunnel syndrome, acute compartment syndromes and peripheral neuropathies, sensory changes, especially those of vibration and thermal perception thresholds, will be detected early and will precede motor loss¹², thereby demonstrating the usefulness of QST as a diagnostic tool. Many clinical sensory tests are valuable in diagnosing carpal tunnel syndrome. This may include the Ten Test, the static and moving two-point discrimination test, von Frey hairs or Semmes-Weinstein monofilament tests, vibrometry, the pressure-specified sensory device¹³; and provocative tests, such as Phalen's wrist flexion test and Tinel's sign.¹⁴ Studies have shown that vibration and light touch sensation perception thresholds will be affected early in carpal tunnel syndrome, and that impaired static and moving two-point discrimination are late findings.¹⁵

Vibration perception threshold (VPT) has been used in the diagnosis of hand-arm vibration syndrome, established by increased VPT in peripheral cutaneous sites.¹⁶ VPT is also used in the assessment of peripheral neuropathy in conditions such as diabetes, stroke, spinal cord injury and age-related sensory loss.^{17, 18} It can identify neuropathy even when patients are asymptomatic.^{17, 18} Vibration threshold testing has also been shown to be useful in assessing symptomatic and asymptomatic alcoholic polyneuropathy¹⁹, neuropathy in cancer patients, peripheral neuropathy and treatment-related neurotoxicity in patients with human immunodeficiency virus.¹¹

b. Prognosis. Abnormalities in sensation have been found to be predictors of poor outcome in several populations. For example, in the acute post-surgery population, the odds of chronic post-surgical pain increase, with an odds ratio of 2.68 for hypoesthesia, and an odds ratio of 6.27 for hyperesthesia.²⁰ Post-operative sensory changes were the *most significant* predictors of chronic post-surgical pain, compared to other factors such as psychological distress and body mass index.

In whiplash disorders, lowered cold pain thresholds and decreased cold pain tolerance are significant predictors of poor recovery.²¹ Significantly, cold hyperalgesia (cold pain threshold > 13°C), results in an increased risk of chronic, severe pain and disability by a very high odds ratio of 26.32 (95% CI: 4.98-139.09).²²

c. Outcome Evaluation. Outcome evaluation is important in determining the progression or stability of a patient with nerve impairment. Outcome measures decrease the subjectivity of sensibility assessment and support clinicians in making objective, valid

and reliable clinical judgements .⁴ This may include determining whether nerve regeneration is occurring and whether treatment is effective. Thus, it will contribute to further clinical treatment planning, delivery of education to the patient and reporting progress to other members of the health care team.²³ Outcome measurement is also critically important in researching the usefulness of treatment techniques, whether surgical, pharmacological, or rehabilitative.²⁴ Jerosch-Herold²³ points out that multiple sensory tests should be used to cover the spectrum of sensory changes that a patient may experience. Ideally, clinicians and researchers should be able to rely on sensation tests which are reliable, valid, responsive over time, standardized, assess clinically meaningful change, and are able to detect a wide range of deficit .²⁴

D. Clinical Measurement

Sensory tests need to be accurate in testing the properties which they propose to test and be able to examine a spectrum of deficits from anaesthesia to normal sensation.²⁴ The best tools are reliable, valid and responsive, and have standardized administration, scoring and interpretation.¹⁰ Ideally, normative data is available.¹⁰ Together, tests that are strong in these properties allow the clinician or researcher to have confidence in the judgments made from the test results.²⁴ Clinical measurement properties have been used in this thesis to evaluate the usefulness of sensory threshold tests. Clear definitions of reliability, validity and responsiveness to change are required for the understanding of clinical measurement properties of sensory tests.

a. Reliability. Reliability is the extent to which a test is free from errors in measurement.²⁵ Mathematically, reliability is the ratio of the variance of true scores to observed variance:²⁶

$$Reliability = \frac{\sigma_{True}^2}{\sigma_{Observed}^2} = \frac{\sigma_{True}^2}{\sigma_{True}^2 + \sigma_{Error}^2}$$

A reliable measure will assess change due to change in health status (true variance) and not due to variations in the assessment technique or tool (error variance).⁵

Reproducibility is the degree that the same outcomes are obtained on repeated testing of an instrument when no change in health status has transpired.²⁷ Reproducibility can be compromised by random measurement error or real within-person variance²⁸, creating what is often referred to as “background noise”.²⁷ To determine reliability, repeated measures are taken, and the differences between scores are used to calculate measurement error.²⁵

Reliability is composed of several, often over-lapping, components: *test-retest*, *intra-rater*, *inter-rater* and *internal consistency*.^{25, 29} *Test-retest reliability* tests the agreement between scores attained on different occasions.²⁹ *Intra-rater reliability* assesses the consistency with which a single observer is able to repeatedly score a test, ideally with identical presentation on each occasion.²⁵ *Inter-rater reliability* is the consistency of measurement between raters when assessing an identical presentation.²⁵ *Internal consistency* measures the homogeneity of scores or items of a test.²⁹

Reliability is measured with a correlation coefficient. Kappa (κ) is used for binary data and represents the amount of agreement corrected for chance.²⁹ The interclass correlation coefficient (ICC) is a relative reliability measure.²⁹ Other measures

can also be used, such as Pearson's correlation coefficient (r). Standard error of measurement is a measure of absolute reliability in the units of the test and is a useful indicator of typical variation between testing occasions.³⁰ Finally, reliability is necessary for validity.

b. Validity. The validity of a test is the extent to which a test measures what it purports to measure.²⁶ Validity is not a property of a test, but the ability to evaluate the intended outcome for a specific purpose and population.²⁹ A gold standard is a recognized reference test and is used to establish criterion validity.²⁹ No measure has been identified as a gold standard for hand sensation assessment.³¹

There are many different types of validity. Three main types are *construct*, *content* and *criterion* validity. *Construct validity* is the ability of a test to measure underlying constructs, such as light touch sensation threshold or pain.²⁵ It can be determined by comparing similarity and divergence with other tests.²⁵ *Content validity* measures whether a test is completely and relevantly assessing the concept which it purports to.²⁵ *Criterion validity* measures how closely a test relates to another test which is considered to be the “gold standard”.²⁶ Although an important measure of validity, there is often a lack of gold standard for testing in health care.²⁶

c. Responsiveness to change. Responsiveness to change, a type of validity, is an important property for tests which are used as outcome measures, as it assesses the degree to which a measure can assess a real change in a clinical state on repeated

testing.²⁷ It provides a more comprehensive understanding of the change in a person's health status. Responsiveness is evaluated using an external criterion to determine whether a person's status has changed³² and the minimum change that is clinically relevant.³² Responsiveness includes the consideration of sensitivity and specificity.

E. Clinical Measurement properties of quantitative sensory tests

Sensation tests salient to this thesis will be explored, namely threshold detection tests, including the Semmes-Weinstein monofilaments test, Weinstein enhanced sensory test, pressure-specified sensory device, Ten Test, current perception threshold, vibration perception threshold, pressure algometry, static and moving two-point discrimination and cold/warm perception tests. Sensory tests not explored in this thesis include functional tests such as the Moberg pick up test and the shape-texture identification test (STI-testTM) and provocative tests such as Phalen's wrist flexion test and Tinel's test.

i. Sensory Threshold tests.

Sensory threshold tests attempt to determine “how much” sensation a person perceives in an affected area. This is determined by a complicated interaction between the peripheral sensory receptors, peripheral nerves and the central nervous system.³³

Methods. Threshold testing typically applies a range of stimuli at standard testing sites to quantify sensory dysfunction.³⁴ Threshold tests identify the minimum stimulus required to produce sensation⁵ or use an intensity rating scale. Two methods commonly used to identify the minimum stimulus identifiable are the method of limits, where an increasing stimulus is applied until sensation is detected, or the method of levels, where

varying intensity of stimuli are applied, and the individual is asked to report whether the sensation was felt (yes/no).³⁴ The method of limits is faster, but results may be affected by reaction time.³⁴ This may be partly ameliorated by slowing ramp speed.³⁴ Threshold values can be compared to age-dependent normative values for various body sites, if available.³⁴ For stimulus intensity rating, a fixed stimulus is applied and the respondent is asked to rate the intensity. Sensation tests suggest that specific nerve receptors and fibres are recruited with various testing modalities; however, research has shown that there are additional sensory signals via application force and vibration through the examiners hand that recruit multiple types of sensory receptors, therefore test selectivity cannot be assumed.³⁵ See Table 2 for a summary of clinical measurements properties of sensation tests.

QST vs. other testing methods. Quantitative sensory testing may be superior to other diagnostic tools, such as computed tomography and magnetic resonance imaging scanning, in the assessment of peripheral nerve conditions.¹¹ QST is superior to nerve conduction velocity (NCV) studies. QST is able to test and quantify the function of small nerve function, whereas NCV testing is limited to assessing large nerve function.¹¹ Also, where NCV assesses the peripheral nervous system, QST assesses the whole pathway.¹¹ Lastly, compared to NCV, QST is easier to perform, is not painful (with the exception of pain threshold testing), and does not require highly-trained persons to perform the test.¹¹

Test order. A recent study has demonstrated that test order in QST is important, as thermal stimulation may cause sensitization to mechanical stimuli, such as pressure algometry.³⁶ Mechanical stimuli does not appear to affect temperature sensitivity.³⁶ Therefore, despite the fact that some standardized protocols may dictate that thermal

testing is to be performed first³⁷, it should, in fact, be performed last in order of QST testing.³⁶

ii. *Light touch/mechanical detection threshold.*

Light touch sensation is mediated by large, myelinated A- β sensory fibres.^{38, 39} Light touch sensation can be evaluated using cotton balls or light touch by a clinician's fingertips to a specific area of innervation by dermatome or peripheral sensory nerve.⁴⁰ The patient is asked to determine whether the sensation differs from a comparative site, such as the contralateral area, or a normally innervated part of the same hand.⁴⁰ This technique does not allow the clinician to quantify responses. A review of light touch testing with this method found inter-rater kappa values to range from 0.31 - 0.90.⁴¹

Monofilament tests: The Semmes-Weinstein monofilament test (SWMT) is a touch threshold test. It measures the lightest pressure that can be detected using test filaments calculated to bend with a specific force (see Table 1).^{23, 40} Protocol suggests that the lightest filament is applied first with three applications and, with the detection of at least one of three applications, the next heaviest filament is tested with the same procedure.⁴² Each test application should be performed with the subject blinded.⁶ Each application should last for 1.0 to 1.5 seconds at the pressure that bends the filament.⁶ The detection threshold is the lightest filament at which at least one of three detections occurred⁴². The Weinstein Enhanced Sensory Test (Bioinstruments Inc., Connecticut) (WEST) is a modification of the SWMT. It uses a testing instrument with five filaments

on one handle.²³ Functional interpretation based on test results have been proposed, but are not substantiated (see Table 1).²⁴

Table 1. Semmes-Weinstein monofilament scale of interpretation

<i>Filament</i>	<i>Interpretation</i>	<i>Force (g)</i>
1.56- 2.83 (Green)*	Normal light touch	0.0045-0.068
3.22- 3.61 (Blue)	Diminished light touch	0.166-0.408
3.84- 4.31 (Purple)	Diminished protective sensation	0.696-2.052
4.56-6.65 (Red)	Loss of protective sensation	3.63-447
Greater than 6.65 (Red-lined)	Untestable (no response)	> 447

*Miniset monofilaments are in bold type.

Note. Adapted from *The Hand: Fundamentals of Therapy* (p. 8), by Judith Boscheinen-Morrin and Bruce Conolly, 2001, Avon, Great Britain: Butterworth-Heinemann. Copyright 2001 by Elsevier Ltd.

Monofilament tests: validity. Few studies have measured the validity of the SWMT or WEST tests. Correlation between SWMT and object recognition for a mixed population was moderate (Spearman’s rank correlation coefficient $r_s = 0.67 - 0.69$).⁴³

Monofilament tests: Reliability. The SWMT is reported to have good inter- and intra-rater reliability.⁴³⁻⁴⁵ Inter-rater reliability for the SWMT has been reported as high (ICC = 0.97, lower 95% CI limit = 0.93),⁴³ Test-retest reliability of the SWMT was reported as high (ICC = 0.84, 95% CI = 0.75 – 0.90) for ten subjects with spinal cord injury and neuropathic pain, and moderate for ten control subjects (ICC = 0.63, 95% CI = 0.45 – 0.76).¹⁷ Test-retest reliability was found to be high in the hand (ICC = 0.87 – 0.99) in 43 children aged six to 12 years old, with threshold values comparable to adults.⁴⁵

Monofilament tests: Responsiveness. The SWMT and WEST tests have been found to have moderate to large effect sizes (ES = 0.73 - 0.80) in populations with median nerve injury and repair.^{23,46} The SWMT was found to statistically show difference in change of touch threshold sensation post nerve repair between three to six versus six to 12 months ($p = 0.0007$, ES 0.73).⁴⁶

Monofilament tests: criticisms. The WEST and SWMT tests have been criticized for having unsatisfactorily wide normative values due to the use of an ordinal scale with increasingly unequal intervals.⁴⁷ The SWMT also takes extensive testing time, decreasing its clinical utility.⁴⁸ The SWMT has also been criticized for requiring skill in application and analysis of findings. It has a potential for the following false positives: 1) Prescribed normative values do not account for variables such as aging and increased skin resistance (callouses);⁴⁰ 2) Increasing or decreasing the filament contact time will alter responses.⁵

Pressure-specified sensory device (PSSD): Another test for light touch sensation threshold is the pressure-specified sensory device, developed by A. L. Dellon in 1991.¹³ This computer-based apparatus consists of two rounded probes attached to a force transducer.⁵ The PSSD can also be used to assess static and moving one- or two-point discrimination.⁴⁹ The touch threshold test delivers increasing pressure through a single applicator prong; the patient presses a trigger when sensation is felt.⁴⁰ Normal values have been reported at 1g/mm² for less than 45 years old and 2.2g/mm² for over 45 years

of age.¹³ Protocol recommends that five trials are performed, with the middle 3 trials averaged.⁴⁰ This device has the benefit of using a continuous scale.⁴⁰

Pressure-specified sensory device: validity: The Wagner Force One Model FDIX 50™ model (Wagner Instruments, Greenwich, Conn) was found to have excellent correlation with force plate measurements (Pearson's $r = 0.99$).⁵⁰ However, using a different algometer (Somedic Algometer Type II, Sweden), applications against a force plate found unacceptable differences (>10%) between algometer and force plate readings.⁵¹

Pressure-specified sensory device: reliability. Research on the reliability of the pressure-specified sensory device for touch threshold is lacking.

Pressure-specified sensory device: responsiveness. PSSD thresholds have been found to be more sensitive than clinical exam in identifying decreased sensory nerve function.⁵² The PSSD was found to have sensitivity (91%) and specificity (82%) in diagnosing carpal tunnel syndrome, equal to nerve conduction studies.¹³ One study has reported favourable sensitivity (100%), but questionable specificity (0% for carpal tunnel syndrome and tarsal tunnel syndrome, 29% for cubital tunnel syndrome and 33% for common peroneal nerve entrapment) in identifying various nerve compression syndromes compared to a “gold standard” test of NCV testing.⁵³

Pressure-specified sensory device: criticisms. There are several drawbacks for this testing method. Skill of use is a potential source of error. Much of the literature about this unit has been published by the developer and therefore needs independent testing.⁴⁰ Reliability and validity testing is limited. The device is handheld by an examiner and is consequently subject to error of fluctuating pressure of application.⁴⁰ Finally, it is also an expensive device.

iii. *Light moving touch*

Light moving touch sensation is mediated by large, myelinated A- β sensory fibres.³⁸ Light moving touch sensation can be evaluated in the same manner as light touch, modified by dragging the contact along the skin surface. Light touch sensory assessment using this technique can quickly define areas of altered sensation.³⁴

Ten Test: The Ten Test was initially described by Strauch et al.³³ as a quick and clinically useful test to quantify sensation loss and to assess for change. It is a quantitative intensity rating assessment of moving light touch sensation. It provides a ratio of sensation lost compared to a client's own normally innervated comparison site.⁴⁸ To conduct the test, a patient reports the level of sensation from one to 10 while light moving touch is simultaneously applied by an examiner's fingertip to the affected palmar fingertip area and an unaffected, similarly innervated area (for example, the index finger of the contralateral hand).³³ Detailed testing protocol can be found at <http://www.youtube.com/watch?v=ktvjsqbIfUM>. If the symptoms are bilateral, then the lip, cheek or bridge of nose is proposed to be used as a reference site.³³ Testing

contraindications include open wounds, or where a normal, similarly innervated reference area is not existing.⁵⁴

Ten Test: validity. Strauch et al.³³ tested for concurrent validity to the Semmes-Weinstein monofilament test. Results produced a Spearman's $r = -0.71$ (95% CI: -0.68 to -1.00, $p < 0.05$). The Ten Test was found to be a valid test in the pediatric population over five years of age.⁵⁵

Ten Test: reliability. Inter-rater reliability in the pediatric population was found to have a kappa agreement of 1.0.⁵⁵ Strauch et al.³³ demonstrated excellent inter-rater reliability of ICC = 0.91 (95% CI: 0.87 - 0.94, $p < 0.05$). Strauch et al.³³ also examined the intra-rater reliability of six raters and found good intra-rater reliability for four out of six raters (ICC = 0.61 – 0.90); however, two raters had ICC values of 0.25 and 0.38.

Ten Test: responsiveness. The Ten Test was found to be able to detect minimal loss of sensation in patients with carpal tunnel syndrome, superior to the WEST, static, and moving two point discrimination tests.⁴⁷ Faught and McKee⁵⁶ determined that using a positivity criterion of scoring < 10 on 3 of 4 of the first to fourth fingers was predictive of carpal tunnel syndrome (sensitivity = 80%, specificity = 48%).

Ten Test: criticisms. The original Ten Test methodology did not allow for assessing hyperesthesia; however, several authors have documented hyperesthesia in several ways. Hyperesthesia has been described by documenting scores $> 10/10$ ⁴⁷, or by

altering the reference, using the 1-10 scale, where 1 = normal with higher numbers representing increasing hyperalgesia⁵⁴; although, it is unclear in this case what the reference for 10/10 on that scale would represent. Another consideration is that a subject may be unaware of bilateral sensory loss, such as in an older population and, therefore, the comparison site may not be normally innervated, as the test assumes. As a subjective test, rating the perception of sensibility may be challenging for some people.⁵⁴ Much of the research has been contributed by the developer and independent testing is needed.

iv. Current perception threshold.

Current perception threshold (CPT) assesses the function of sensory nerves with the use of electricity.^{57, 58} It is measured by applying 1cm diameter, gold-plated, gel-coated electrodes⁵⁸ to a glabrous skin area representing a cutaneous nerve or dermatomal distribution. A microprocessor-controlled electrical stimulator delivers a sinusoidal constant alternating current at one of three different frequencies: 5Hz, 250Hz or 2000Hz.⁵⁹ The constant current delivery compensates for variations in skin resistance, skin thickness and moisture.⁵⁸ CPT provides a minimal stimulus using an automated protocol that can be administered following little training.⁶⁰ Subjects control the delivery of stimulation and respond at the moment that they feel the onset of sensation. The CPT is the point at which sensation is initially perceived. Normal range values are 6 to 13 (no units), values less than 6 are hypersensitive, and values higher than 13 demonstrate a loss of sensation.⁶¹ CPT stimulates the nerve fibres directly; the intensities delivered are below the stimulation of skin receptors.⁵⁹ The three different frequencies assess three types of afferent neurons: large myelinated A- β fibres (2000Hz), A- δ fibres (250Hz) and

C fibres (5Hz).⁶² Comparisons with other types of study have supported the conclusion of nerve selectivity based on frequency.⁵⁹ This includes nerve conduction studies, thermal sensory testing, vibration perception threshold, somatosensory evoked potential studies, histopathological and pharmacological studies.⁵⁹

Current perception threshold: validity. Lowenstein and colleagues⁶² investigated the relationship between CPT and quantitative sensory testing (QST). They tested 27 healthy women. The authors found that CPT testing at 5Hz had moderate correlation with thermal testing (Spearman's $r = 0.49$, $p = 0.01$), CPT at 2000Hz had moderate correlation with vibration perception threshold (Spearman's $r = 0.50$, $p = 0.01$), but cold thermal testing did not correlate with CPT testing at 250Hz ($p = 0.30$) or 5Hz ($p = 0.10$).⁶² Previous research has demonstrated moderate correlation between CPT testing at 250Hz with warm sensation threshold testing (Spearman correlation coefficient, $r = 0.46$, $p < 0.005$), and testing at 2000Hz with vibration sensation (Spearman correlation coefficient, $r_s = 0.42$, $p < 0.005$).⁶³ Katims et al.⁵⁷ found high correlation ($r = 0.79$, $p < 0.001$) between CPT and nerve conduction velocity studies assessing the severity of compression neuropathy in people with diabetes.

Current perception threshold: reliability. Lowenstein, Jesse & Kenton⁶² found a significant difference in repeatability when tested one week apart ($p > 0.05$) on 27 healthy women. Park, Wallace and Shulteis⁶⁴ found poor test-retest reliability in 19 healthy subjects for sensory threshold testing (mean difference score: CPT 5HZ threshold = $5.24 \pm 229\%$; CPT 250Hz threshold = $3.74 \pm 1.21\%$; CPT 2000Hz threshold = $0.98 \pm$

188.32%). Katims et al.⁵⁷ found coefficients of variation (6.5% at 2000Hz, 13.7% at 250Hz and 27.5% at 5Hz) on repeated CPT testing of the median and ulnar nerves in seven diabetic patients on hemodialysis.

Current perception threshold: responsiveness. No data was discovered that assessed the responsiveness of CPT in hand sensory function.

Current perception threshold: criticisms. CPT cannot localize a deficiency within the peripheral or central nervous system.⁵⁹ As the above summations reported, repeatability is poor and validity is limited. The equipment is also expensive.

v. *Vibration.*

Vibration is mediated by large, myelinated A- β nerve fibres.³⁸ Increased vibration perception threshold (VPT) is one of the first signs of peripheral neuropathies and nerve entrapment¹², and of reinnervation following nerve repair.⁹ VPT testing has been shown to be valuable as a non-invasive diagnostic technique in evaluating nerve compression, acute compartment syndromes, peripheral neuropathies, and nerve repair.^{4, 12}

Tuning forks: Tuning forks are used by applying vibration through the tuning fork and comparing the sensation felt at a second site for similarity.⁴⁰ The traditional use of tuning forks to assess vibration sensibility is of little clinical value, as they are not quantifiable and have inconsistent inter-rater and intra-rater reliability.⁴ Tuning forks have been criticized for having large variations in application force and oscillation

frequency, having uncontrolled force of application, and being influenced by the vibration of the hand of the examiner holding the instrument.³⁵

Vibration threshold testing: Vibrometers are superior to tuning forks in assessing vibration sensation. Vibrometers assess vibration perception threshold (VPT) via a computer controlled device. This device delivers vibration through an applicator held by an examiner against a person's skin surface. The area being tested should be resting in a comfortable position. Commonly, VPT assessment is assessed at the palmar fingertip surface; however, this area is prone to callouses, and it is not known to what degree this affects measurement.¹⁶ Age has been disputed as having an effect on VPT thresholds in the hand, as it does for the foot, and may need to be accounted for in interpreting results, although height and gender consistently have not impacted VPT in the hand, as they may for the foot.^{16, 65} The delivery of vibration is standardized. The frequency, intensity and ramp speed can be controlled⁴⁰, as well as other parameters such as ramp speed. Slowing ramp speed may help to decrease the error involved with response time delay. Differences occur in the vibration perception thresholds of glabrous (non-hairy) and non-glabrous (hairy) skin, with greater discriminative power of glabrous skin.⁶⁶ Therefore, VPT should be assessed at glabrous skin sites. There are several testing methods to determine VPT. In the method of limits, for example, an increasing intensity of vibration is delivered through the applicator until the person senses the vibration, at which time a button is pressed to indicate the first onset of sensation. Repeated trials are performed. Decreased reliability in the first trial has shown the importance of allowing a practice trial.¹²

Vibration: validity. Vibration perception threshold has been shown to increase with increasing severity of diabetic neuropathy.⁶⁷ VPT did not discriminate between diabetics asymptomatic for neuropathy and their age-matched controls, as well as warm thermal threshold testing did.⁶⁷ In other words, warm thermal testing was more sensitive than VPT in detecting early, asymptomatic neuropathy in people with diabetes.

Vibration: reliability. Vibrometers have been shown to reliably assess VPT when testing protocol for individual units are followed.⁶⁸ Inter-rater reliability was found to have an ICC = 0.56 (right hand) and 0.88 (left hand) for two observers assessing 39 healthy people.⁶⁵ The inter-rater reliability of the Vibratron II was substantial (ICC = 0.98, lower 95%CI = 0.97).⁴³

Peters et al.⁶⁵ found ICC to range from 0.55 - 0.80 for short-term repeated testing (15 minutes), and ICC = 0.77 - 0.95 for repeated testing after 24 hours in the hand. Lowenstein, Jesse & Kenton⁶² found no significant difference in repeatability when repeat testing was performed one week apart ($p = 0.30$) on 27 healthy women. Similarly, van Deursen et al.,⁶⁹ found overall excellent test-retest reliability for the Biothesiometer (Bio-Medical Instrument Co., Newbury, OH, USA) for both replication to replication and day to day retest reliability, for both controls and in diabetic patients with peripheral neuropathy. Mahbub et al.⁷⁰ found that glabrous skin had better test-retest repeatability than non-glabrous skin, with excellent repeatability testing at glabrous finger sites in healthy subjects (ICC = 0.84 to 0.91 for with-in session testing, and high ICC for intersession testing for 7/8 subjects). Grunert et al.¹² reported test-retest reliability of ICC

= 0.76 - 0.87 for within session repeatability. Test-retest reliability was reported as substantial (ICC = 0.90, 95%CI = 0.84 – 0.94) for ten subjects with spinal cord injury and neuropathic pain, and substantial (ICC = 0.86, 95% CI = 0.79 – 0.91) for ten control subjects tested approximately 3 weeks apart.¹⁷ Overall, VPT has consistently demonstrated excellent to substantial test-retest reliability.

Vibration: responsiveness. Data on the responsiveness of VPT is lacking.

Vibration: criticisms. One concern with vibrometers is the issue of reaction time, with the consequence of an overestimation of sensory threshold.⁶⁸ Reaction time is influenced by concentration, drowsiness, and boredom.⁶⁸ A learning effect has been demonstrated by several authors, demonstrating the importance of a trial application.¹² Another limitation is that VPT assesses the whole sensory system and cannot localize a deficit to the peripheral or central nervous system.^{16, 68} A final limitation to note with VPT is that, being subjective, it is not reliable in distinguishing between organic or psychogenic causes.⁶⁸ Although normative data has been published, they are only useful if the exact protocol, such as equipment used and testing site is replicated.³⁹ Vibrometers are expensive (\$1 000 – \$30 000) and not widely used.²⁴

Overall, VPT appears to be a valid and reliable tool for assessing nerve dysfunction and in assessing a change in health status. It is especially useful in assessing early stages of sensory loss. However, due to the high cost, it may remain more commonly utilized for research, as opposed to clinical, purposes.

vi. Pain.

Sharp/dull: Pain is mediated by A- δ and C fibres. Sharp/dull testing has been traditionally performed by using the pinprick method with pins or small diameter probes, and the individual is asked to determine if the sensation is sharp or dull.³⁴ Sharp sensation thresholds have been shown to closely correlate with pain thresholds.³⁴

Sharp/dull: reliability. Sharp/dull testing with the pinprick method was found to have slight to substantial agreement between two raters, dependent on the nerve root tested ($\kappa = 0.16 - 0.67$).⁷¹

Pressure algometry: Pressure algometry can be used to assess pressure-pain threshold, the point at which an applied force is perceived as pain. Pressure algometry is the most commonly used test for mechanical allodynia in deep tissues³⁴ and is used clinically to determine trigger point tenderness in myofascial pain syndromes.⁷² Pressure-pain thresholds allow the quantification of tenderness that may be difficult to measure with other methods, and allows assessment of change.⁵⁰ The devices are typically hand-held, with a 1cm² application surface area, reports force in kilograms of force,⁵⁰ and should be applied at the rate of 1 kg/second.⁷³

Pressure algometry: validity. Two studies measuring the correlation with force plate measurements were described for pressure algometry under the above section for touch threshold testing.

Pressure algometry: reliability. Group test-retest reliability was found to range from moderate to good (ICC = 0.78 - 0.93) depending on the site tested using a hand-held pressure algometer in 20 women with neck pain; however, the authors found considerable individual variation, indicating that pressure algometry may be useful for research purposes, but less reliable in a clinical setting.⁷⁴ Nussbaum⁷⁵ also found high group test-retest reliability (ICC = 0.93 – 0.98). A learning effect was observed, hence, allowing an initial trial for learning purposes improved reliability.⁷⁵ Another study found test-retest reliability to be moderate for 15 participants with temporomandibular disorders (ICC = 0.63).⁷² Inter-rater reliability was found to be good between two raters (ICC = 0.74 - 0.89).⁷⁵

Pressure algometry: criticisms. Analyzing pressure algometry in measuring pain threshold has been challenged by methodological issues. A lack of documentation of rate of pressure application, lack of training in the equipment and procedure, habituation after several applications, and participation of the patient.⁵¹ Additionally, the use of a verbal response from the patient to indicate threshold depends on reaction time from the examiner.⁵¹ Overall, there is limited research published for this technique.

vii. *Static/moving 2 point discrimination.*

Two-point discrimination measures the innervation density of cutaneous mechanoreceptors. In the fingertips, there is high innervation density, producing small receptive fields, therefore, allowing highly detailed assessment of sensory input.

Static and moving two-point discrimination test: The static two-point discrimination test is a widely used measure of innervation density testing, although it lacks validity and reliability.²³ This test assesses the smallest distance between two points at which a person can discriminate between one and two points.²⁴ Normal moving two-point discrimination has been defined as 2mm at the distal fingertip.⁹ Originally performed using the end points of a paperclip⁹, standardized equipment has now been developed. A commonly used device for the measurement of two-point discrimination is the Dellon-Mackinnon Disk-Criminator™ (Dellon-Mackinnon Disk-Criminator™, P. O. Box 16392, Baltimore, Maryland). This disc has a range of calipers, spaced from 2 to 15 mm apart which are applied perpendicularly to the finger pad, although it is often incorrectly applied along the long axis.⁵ Testing is performed with the subject's vision occluded.⁶ The test protocol defined by Jerosch-Herold⁷⁶ is performed by using the widest calliper distance first, applying ten applications of one or two points. If the subject correctly differentiates at least seven of the applications, then the next calliper is used in the same manner. This test can also be performed as a moving two-point discrimination test, whereby the same protocol as above is employed, with the variation that the callipers are moved along the skin surface.⁹

Two-point discrimination test: validity. The validity of the two-point discrimination as a test for innervation density or spatial threshold has been questioned.²⁴ Moving two-point discrimination was previously proposed to be an indicator of hand function; however, the test has poor correlation with functional hand tests such as the

pick-up test, shape and texture identification when corrected for age, delay time of surgery post injury (sutured nerves post complete division) and follow-up time ($r = -.05 - 0.13$).⁷⁷ Construct validity has not been assessed.²⁴ Concerns about the validity of this test include the ability to consistently apply equal force of application and the interference of application related clues.^{24,35} Interference of vibratory stimuli was found to emanate from examiners at amplitudes sufficient to stimulate sensory receptors, and the interference was even more pronounced if the examiner supported the part being tested with their other hand.³⁵ Bell-Krotoski & Buford³⁵ also found that force differed between the applications of one or two points. Finally, the two-point discrimination test is a late finding in sensory loss.³¹

Two-point discrimination test: reliability. Inter-rater reliability has been found to be high for moving (ICC: 0.99, lower 95%CI = 0.98) and static two point discrimination (ICC: 0.99, lower 95%CI = 0.98) in a mixed population.⁴³ The 30 study subjects had either vision-impairment (n = 14), normal sensation (n = 6) and previous nerve impairment (n = 8) (Novak, et al. 1993). The overarching criticism was that the application pressure could not be controlled⁴² and that intra-rater reliability was found to be poor.³⁵

Two-point discrimination test: responsiveness. Several studies have found that two-point discrimination has poor test responsiveness, with effect sizes of 0.00 - 0.11^{23,46,78}. This was most noteworthy in populations of people with complete nerve transections where it is extraordinary to find values below 15mm (the upper limit of several testing

devices).²⁴ Novak⁵ points out that two-point discrimination wider than the distal finger pad is, functionally, no two-point discrimination. Two-point discrimination did not improve when patients with post nerve repair were tested at three to six versus six to 12 months (mean difference = 0).⁴⁶ Other sensory and motor tests improved; therefore, two-point discrimination was not an informative indicator of change in sensory function.⁴⁶ Other studies found that the two-point discrimination test had strong flooring effects.⁷⁸ An error of one to two mm on testing should be assumed.³¹

Two-point discrimination test: criticisms. Two-point discrimination tests are not useful in compression syndromes; as abnormal findings will not occur until late in the syndrome.⁶ The lack of reliability, validity and test responsiveness indicate that other sensory tests would be more usefully employed.

viii. Temperature.

Cold sensation is believed to be mediated by small myelinated nerve fibres (A δ) and warm sensation by unmyelinated warm specific C nerve fibres.¹¹

Thermal threshold testing: Thermal threshold testing can be quantified using a computer-controlled device such as the TSA-II (Medoc Advanced Medical Systems, Minneapolis). Temperature rises or falls via a thermode applied to a patient's skin. The person tested presses a button when either warm or cold sensation is initially felt. Quantified thermal testing has been found to be useful in diabetic polyneuropathy, where

it is able to identify sensation loss early, even before symptoms or abnormalities on nerve conduction studies appear.⁶⁷

Thermal threshold testing: validity. Warm and cold perception thresholds have been shown to increase with increasing severity of diabetic neuropathy.⁶⁷ Thermal thresholds were not shown to correlate with subjective neuropathic pain symptoms in people with spinal cord injury.¹⁷

Thermal threshold testing: reliability. A systematic review of 21 studies of thermal detection threshold reliability found several main results: 1) the method of limits and method of levels demonstrated similar reproducibility; 2) the studies reviewed reported a wide range of test-retest reliability values for cold and warm detection thresholds, from poor to excellent reliability; the most commonly reported was fair reliability; 3) the populations studied were primarily control populations and people with diabetes.⁷⁹

Thermal threshold testing: responsiveness. Shukla et al.,⁸⁰ found that thermal threshold testing was more sensitive to identifying small nerve fibre neuropathy than nerve conduction velocity tests and physical examination in 25 people with suspected small fibre neuropathy.

Thermal threshold testing: criticisms. Several concerns limit the usefulness of thermal perception testing. Cold pain perception normative values are broad.³⁴ There are also concerns about reproducibility of results, and limited studies on validity exist. Few different populations have been studied.

Table 2. Summary of clinical measurement properties of sensory assessments.						
Modality	Sensory test	Sensory nerve fibres	Population	Reliability	Validity	Responsiveness
Light touch	Semmes-Weinstein Mono-filament test (SWMT)	Large, myelinated A-β	Varied adult groups: control, visually impaired adults, adults with previous nerve injury.	Inter-rater reliability high (ICC = 0.97, lower 95%CI limit = 0.93) ⁴³	Concurrent validity: b/w SWMT and object recognition: (Spearman's $r = 0.67-0.69$). ⁴³	
			Adults with spinal cord injury	Test-retest reliability high (ICC = 0.84, 95%CI = 0.75 – 0.90). ¹⁷		
			Adult control population	Test-retest reliability moderate (ICC = 0.63, 95%CI = 0.45 – 0.76). ¹⁷		
			Pediatric index finger	Test-retest reliability high (ICC = 0.87 – 0.99). ⁴⁵		
	SWMT and Weinstein enhanced sensory test (WEST)		Adults with median nerve injury and repair			Moderate to large effect sizes (ES = 0.73 - 0.80). ^{23, 46}
	Pressure-specified sensory device (PSSD)		No population (force plate measurement)		Wagner Force One Model FDIX 50™ model excellent correlation with force plate measurements (Pearson's $r = 0.99$). ⁵¹	

			No population (force plate measurement)		Somedic Algometer Type II: applications against a force plate found unacceptable differences (>10%). ⁵²
			Adults suspected of having carpal tunnel syndrome		Sensitivity (91%); specificity (82%) in diagnosing carpal tunnel syndrome, equal to NCV studies. ¹³
			Adults with symptoms of chronic nerve compression.		Favourable sensitivity (100%), but poor specificity (0% for carpal tunnel syndrome and tarsal tunnel syndrome, 29% for cubital tunnel syndrome and 33% for common peroneal nerve entrapment) in identifying various nerve compression syndromes compared to a “gold standard” test of NCV testing. ⁵³
Light moving touch	Ten Test	Large, myelinated A-β	Adults with peripheral nerve disorders	Interrater reliability between two raters, and found an interclass correlation coefficient of ICC = 0.91 (95% CI: 0.87 - 0.94, $p < 0.05$). ³³	Concurrent validity to the SWMT: Spearman’s $r = -0.71$ (95% CI: -0.68 to -1.00, $p < 0.05$). ³³
			Children (age 1 - 12 years)	Good intrarater reliability for 4/6 raters (ICC = 0.61 – 0.90); however, two raters had ICC values of 0.25 and 0.38. ³³ Interrater reliability in the pediatric population had a kappa agreement of 1.0. ⁵⁵	Validated in the pediatric population over five years of age. ⁵⁵

			29 adults with diabetes	Test-retest reliability: repeated CPT testing of the median and ulnar nerves of 7 of the participants resulted in coefficients of variation of (6.5% at 2000Hz, 13.7% at 250Hz and 27.5% at 5Hz). ^{57\}	High correlation ($r = 0.79$, $p < 0.001$) between CPT and nerve conduction velocity studies. ⁵⁷
Vibration	Vibrometers	Large, myelinated A-β	Adults with diabetic neuropathy		VPT increases with increasing severity of diabetic neuropathy. ⁶⁷
			39 healthy adults	Inter-rater reliability: ICC = 0.56 (right hand) and 0.88 (left hand) for two observers assessing 39 healthy people. ⁶⁵	
			Varied adult groups: control, visually impaired, or with previous nerve injury.	Inter-rater reliability of the Vibratron II: was substantial (ICC = 0.98, lower 95%CI = 0.9767). ⁴³	
			39 healthy adults	Test-retest reliability: ICC ranged from 0.55 - 0.80 for short-term repeated testing (15 minutes), and ICC = 0.77 - 0.95 for repeated testing after 24 hours in the hand. ⁶⁵	
			27 healthy women	Test-retest reliability: no significant difference in repeatability when repeat testing was performed one week apart ($p = 0.30$). ⁶²	

		<p>Healthy adults and diabetic patients with peripheral neuropathy</p> <p>Healthy adults</p>	<p>Test-retest reliability: excellent for the Biothesiometer for both replication to replication and day to day retest reliability.⁶⁹</p> <p>Test-retest reliability: excellent: (custom equipment) ICC = 0.84 to 0.91 for with-in session testing, and high ICC for intersession testing for 7/8 subjects.⁷⁰</p>	
		<p>10 subjects with spinal cord injury and neuropathic pain and 10 control subjects.</p>	<p>Test-retest reliability: ICC = 0.76 - 0.87 for within session repeatability.¹²</p> <p>Test-retest reliability: high (ICC = 0.90, 95%CI = 0.84 – 0.94) for 10 subjects with SCI and neuropathic pain, and substantial (ICC = 0.86, 95% CI = 0.79 – 0.91) for 10 control subjects.</p>	
Pain	Pinprick	A-δ nerve and C nerve fibres		Slight to substantial agreement between two raters, dependent on the nerve root tested ($\kappa = 0.16 - 0.67$). ⁷¹
	Pressure algometry	A-δ and C nerve fibres	20 women with neck pain	<p>Group test-retest reliability: ranged from moderate to good (ICC = 0.78 - 0.93) depending on the site tested; however, there was considerable individual variation, suggesting that pressure algometry may be useful for research purposes, but less reliable in a clinical setting.⁷⁴</p> <p>See under the light touch section</p>

		35 healthy subjects	Group test-retest reliability high (ICC = 0.93 – 0.98). ⁷⁵
		15 participants with temporomandibular disorders	Test-retest reliability: moderate for (ICC = 0.63). ⁷²
		35 healthy subjects	Interrater reliability: good between two raters (ICC = 0.74 - 0.89). ⁷⁵
Static/ moving 2 point discrimination	Disk-Criminator ^T ^M , PSSD, calipers	Large, myelinated A-β	Poor correlation with functional hand tests such as the pick-up test, shape and texture identification ($r = -.05 - 0.13$). ⁷⁷
		Varied adult groups: control, visually impaired adults, adults with previous nerve injury. Adults with median nerve injury and repair	Inter-rater reliability: high for moving (ICC: 0.99, lower 95%CI = 0.98) and static two point discrimination (ICC: 0.99, lower 95%CI = 0.98). ⁴³
			Poor test responsiveness, with effect sizes of 0.00 - 0.11. ^{23, 46, 78}
Temperature	Computer-controlled thermodes	Cold sensation: small myelinated nerve fibres (Aδ); warm sensation: C fibres	Warm and cold perception thresholds have been shown to increase with increasing severity of diabetic neuropathy. ⁶⁷

	<p>The populations studied in this systematic review were primarily control populations and people with diabetes</p>	<p>A systematic review of 21 studies of thermal detection threshold reliability found: 1) the method of limits and method of levels demonstrated similar reproducibility; 2) test-retest reliability: a wide range of values reported for cold and warm detection thresholds, from poor to excellent reliability; the most commonly reported was <u>fair</u> reliability.</p>
	<p>25 adults with suspected small fibre neuropathy</p>	<p>Thermal threshold testing was more sensitive to identifying small nerve fibre neuropathy than NCV tests and physical examination.⁸⁰</p>

ICC = interclass correlation coefficient

CI= confidence interval

NCV = nerve conduction velocity tests

LR = likelihood ratio

F. Challenges in the use of sensation assessments

a. Subjective nature of clinical tests. QST are psychophysical tests.⁸⁰ They are subjective and very sensitive to subtle changes in methodology, such as testing site, pressure of stimulator application, stimulator size and training of subjects.¹¹ The client must be alert, compliant and able to follow instructions.⁶⁸ In fact, responses can only be considered to be valid if the patient is cooperative, and, although testing procedures can be used to show inconsistencies in responses, there is no decisive way to determine whether a client is responding truthfully.¹¹

b. Lack of utilization by clinicians. There is a dearth of information on the usage of sensation testing by clinicians. A single example is a recent survey of 381 clinicians from a variety of clinical backgrounds. Fifty three percent used quantitative sensory testing “sometimes” or routinely in practice as an outcome measure for people with neck pain.⁸¹ This seems a high estimate. The authors of the study suggest that the interpretation of QST may have been a factor.⁸¹

G. Summary

A considerable amount of further research needs to be done in the field of sensory testing to determine the most useful sensory testing methods for both research and clinical purposes. Sensory threshold testing would ideally combine several measures, including threshold testing and intensity rating. It needs to be remembered that sensory testing is imperfect, and issues with application techniques will stimulate several sensory

receptors.³⁵ Isolation of stimulus/receptor cannot be assumed.³⁵ For research purposes, the vibration threshold test appears to have the strongest clinical measurement properties. However, in most clinical settings VPT will not be available due to the high cost of equipment.

In the clinical setting, as previously described, tools need to be reliable, valid, responsive over time, assess clinically meaningful change, standardized, able to detect a wide range of deficit and able to assess the properties of the nervous system for which they are intended.²⁴ They also need to be reasonably inexpensive, practical and interpretable without extensive training.

This review has found that the test which may best meet these requirements is the Ten Test. The Ten Test has reported good inter-rater reliability, validity and responsiveness; the test is standardized and assesses loss of sensation across a wide spectrum of sensibility deficit, with the exception of hyperalgesia, although modifications have been proposed to accommodate for this. It is clinically appealing as it is a quick test and requires no instruments. The Ten Test does suffer from a lack of clinical measurement testing. Also, much of the current literature requires confirmation through further research with improved methodologies on larger groups.

Purpose

a. Purpose. The Ten Test is an efficient test which requires no special equipment to perform. If it also demonstrates good clinical measurement properties, then this test may help to increase the utilization of sensibility assessment by clinicians. The main purpose of this research is to provide estimation of the test-retest reliability and concurrent validity of the Ten Test. Concurrent validity has been assessed compared to the SWMT, but current perception threshold and vibration perception threshold tests may be considered more rigorous comparison tests, being applied in a standardized manner by computer controlled devices. Further, CPT is a 25 point ratio scale and VPT is a ratio scales. Test-retest reliability will be assessed from repeated trials within, approximately, a one hour session for the Ten Test, as well as for CPT and VPT. Test-retest reliabilities will provide information on the comparative repeatability of each of the tests, as well as absolute reliabilities (MDC) for use in clinical practice. A visual version of the Ten Test was created with the intention of providing an alternate version which may improve the user-friendliness of the tool, and potentially improve the accuracy of reporting.

The specific research questions being studied are:

i. What are the same-day test-retest reliabilities (ICC, SEM and MDC_{90}) for the Ten Test, vibration perception threshold and current perception threshold in the population of people with decreased sensation in their hands (any diagnosis)?

ii. What is the concurrent validity (Spearman's r_s) for the Ten Test, compared to vibration perception threshold and the current perception threshold tests in the population of people with decreased sensation in their hands (any diagnosis)?

iii. Does a modified version of the Ten Test, in a visual format, correlate to the original version? Which version is preferred by participants?

b. Overview of the thesis format. This thesis will be presented in manuscript style. Chapter One is an overview of current knowledge of sensation threshold testing, as well as a review of reliability and validity. Chapter Two will provide two manuscripts to be submitted to journals for publication. Chapter Three will be a discussion on the future directions for research in this field.

References

1. The Institute of Musculoskeletal Health and Arthritis. Facts & figures. May 14, 2013; <http://www.cihr-irsc.gc.ca/e/27028.html>. Accessed August 8, 2013.
2. Public Health Agency of Canada. Life with arthritis in Canada: A personal and public health challenge. July 19, 2010; <http://www.phac-aspc.gc.ca/cd-mc/arthritis-arthrite/lwaic-vaaac-10/8-eng.php>. Accessed August 8, 2013.
3. Public Health Agency of Canada. Cardiovascular disease - Economic burden of illness. April 22, 2013; http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/cvd_ebic-mcv_femc-eng.php. Accessed August 8, 2013.
4. Bear-Lehman J, Abreau B. Evaluating the hand: issues in reliability and validity. *Physical Therapy*. 1989;69(12):11.
5. Novak CB. Evaluation of hand sensibility: A review. *Journal of Hand Therapy*. 2001;14(4):266-272.
6. Boscheinen-Morrin J, Conolly B. Assessment. *The Hand: Fundamentals of Therapy*. Vol 3rd: Butterworth-Heinemann; 2001:1-12.
7. Lundborg G. Brain plasticity and hand surgery: an overview. *Journal of Hand Surgery (British and European Volume)*. 2000;25(3):242-252.
8. Merzenich MM, Jenkins WM. Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience. *Journal of Hand Therapy*. 1993;6(2):89-104.
9. Dellon A. *Evaluation of sensibility and re-education of sensation in the hand*. Baltimore, Md: Williams & Wilkins; 1981.
10. Warwick D, Dunn R, Melikyan E, Vadher J. Assessment. *Hand Surgery: Oxford Specialist Handbooks in Surgery*; 2009:1-37.
11. Zaslansky R, Yarnitsky D. Clinical applications of quantitative sensory testing (QST). *Journal of the Neurological Sciences*. 1998;153(2):215-238.
12. Grunert BK, Wertsch JJ, Matloub HS, McCallum-Burke S. Reliability of sensory threshold measurement using a digital vibrogram. *Journal of Occupational Medicine*. 1990;32(2):100-102.
13. Weber RA, Schuchmann JA, Albers JH, Ortiz J. A prospective blinded evaluation of nerve conduction velocity versus pressure-specified sensory testing in carpal tunnel syndrome. *Annals of plastic surgery*. 2000;45(3):252-257.
14. Faught B. *Efficacy of clinical tests in the diagnosis of carpal tunnel syndrome*. Toronto: Graduate Department of Exercises Sciences, University of Toronto; 2001.
15. Gelberman RH, Szabo RM, Williamson RV, Dimick MP. Sensibility testing in peripheral-nerve compression syndromes. An experimental study in humans. *The Journal of Bone & Joint Surgery*. 1983;65(5):632-638.
16. Gerr F, Letz R. Vibrotactile threshold testing in occupational health: A review of current issues and limitations. *Environmental Research*. 1993;60(1):145-159.
17. Felix ER, Widerstrom-Noga EG. Reliability and validity of quantitative sensory testing in persons with spinal cord injury and neuropathic pain. *Journal of Rehabilitation Research and Development*. 2009;46(1):69-83.

18. Liu W, Lipsitz LA, Montero-Odasso M, Bean J, Kerrigan DC, Collins JJ. Noise-enhanced vibrotactile sensitivity in older adults, patients with stroke, and patients with diabetic neuropathy. *Archives of Physical Medicine and Rehabilitation*. 2002;83(2):171-176.
19. Sosenko JM, Soto R, Aronson J, Kato M, Caralis PV, Ayyar DR. The prevalence and extent of vibration sensitivity impairment in men with chronic ethanol abuse. *Journal of studies on alcohol*. 1991;52(4):374-376.
20. Johansen A, Romundstad L, Nielsen CS, Schirmer H, Stubhaug A. Persistent postsurgical pain in a general population: Prevalence and predictors in the Tromsø study. *Pain*. 2012;153(7):1390-1396.
21. Sterling M. Does knowledge of predictors of recovery and nonrecovery assist outcomes after whiplash injury? *Spine*. 2011;36(25S):S257-S262.
22. Sterling M, Hendrikz J, Kenardy J. Similar factors predict disability and posttraumatic stress disorder trajectories after whiplash injury. *Pain*. Vol 152: Elsevier; 2011:1272-1278.
23. Jerosch-Herold C. A study of the relative responsiveness of five sensibility tests for assessment of recovery after median nerve injury and repair. *The Journal of Hand Surgery: British & European Volume*. 2003;28(3):255-260.
24. Jerosch-Herold C. Assessment of sensibility after nerve injury and repair: a systematic review of evidence for validity, reliability and responsiveness of tests. *The Journal of Hand Surgery: British & European Volume*. 6 2005;30(3):252-264.
25. Domholdt E. *Physical Therapy Research; Principles and Applications*. Philadelphia, Pennsylvania: W.B.Saunders Company; 2000.
26. Streiner DL, Norman GR. "Precision" and "Accuracy": two terms that are neither. *Journal of clinical epidemiology*. 2006;59:327-330.
27. de Vet HCW, Bouter LM, Bezemer PD, Beurskens AJHM. Reproducibility and responsiveness of evaluative outcome measures. *International Journal of Technology Assessment in Health Care*. 2001;17(4):479-487.
28. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *Journal of chronic diseases*. 1985;38(1):27-36.
29. MacDermid J, Michlovitz, S. Incorporating outcome measures into evidence based practice. In: Law M, MacDermid, J., ed. *Evidence-based Rehabilitation*. 2nd ed. Thorofare, NJ: Slack Incorporated; 2008.
30. Lexell JD, D. Y. . How to assess the reliability of measurements of rehabilitation. *American Journal of Physical Medicine & Rehabilitation*. 2005;84(9):4.
31. Moberg E. The unsolved problem - How to test the functional value of hand sensibility. *Journal of Hand Therapy*. 1991;4(3):105-110.
32. Stratford PW, Binkley, J., Riddle, D. Health status measures: strategies and analytic methods for assessing change scores. *Physical Therapy*. 1996;76(10).
33. Strauch B, Lang A, Ferder M, Keyes-Ford M, Freeman K, Newstein D. The Ten test. *Plastic and Reconstructive Surgery*. 1997;99(4):1074-1078.

34. Walk D, Sehgal N, Moeller-Bertram T, et al. Quantitative sensory testing and mapping: A review of nonautomated quantitative methods for examination of the patient with neuropathic pain. *The Clinical Journal of Pain*. 2009;25(7):632-640.
35. Bell-Krotoski JA, Buford WL. The force/time relationship of clinically used sensory testing instruments. *Journal of Hand Therapy*. 1997;10(4):297-309.
36. Gröne E, Crispin A, Fleckenstein J, Irnich D, Treede R-D, Lang PM. Test order of quantitative sensory testing facilitates mechanical hyperalgesia in healthy volunteers. *Journal of Pain*. 2012;13(1):73-80.
37. Rolke R, Baron R, Maier C, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): Standardized protocol and reference values. *Pain*. 2006;123(3):231-243.
38. Vinik A, Suwanwalaikorn S, Stansberry K, Holland M, McNitt P, Colen L. Quantitative measurement of cutaneous perception in diabetic neuropathy. *Muscle & Nerve*. 1995;18(6):574-584.
39. Yarnitsky D. Quantitative sensory testing. *Muscle and Nerve*. 1997;20(2):198-204.
40. MacDermid JC, Doherty T. Clinical and electrodiagnostic testing of carpal tunnel syndrome: a narrative review. *The Journal of Orthopaedic and Sports Physical Therapy*. 2004;34(10):565-588.
41. Cook C, Hegedus E. Physical examination tests for neurological testing and screening. *Orthopedic Physical Examination Tests; An Evidence-Based Approach*. Vol 1st: Pearson Education; 2008:31-33.
42. Bell-Krotoski J, Weinstein S, Weinstein C. Testing sensibility, including touch-pressure, two-point discrimination, point localization, and vibration. *Journal of Hand Therapy*. 1993;6(2):114-123.
43. Novak C, Mackinnon S, Williams J, Kelly L. Establishment of reliability in the evaluation of hand sensibility. *Plastic and reconstructive surgery*. 1993;92(2):311-322.
44. Bell-Krotoski J, Tomancik E. The repeatability of testing with Semmes-Weinstein monofilaments. *The Journal of hand surgery*. 1987;12(1):6.
45. Thibault A, Forget R, Lambert J. Evaluation of cutaneous and proprioceptive sensation in children: a reliability study. *Developmental Medicine and Child Neurology*. 1994;36(9):796-812.
46. Rosen B, Dahlin LB, Lundborg G. Assessment of functional outcome after nerve repair in a longitudinal cohort *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*. 2000;34(1):71-78.
47. Patel MR, Bassini L. A comparison of five tests for determining hand sensibility. *Journal of Reconstructive Microsurgery*. 1999;15(7):4.
48. Strauch B, Lang A. The Ten test revisited. *Plastic and Reconstructive Surgery*. 2003;112(2):593-594.
49. Dellon ES, Mourey R, Dellon A. Human pressure perception values for constant and moving one and two-point discrimination. *Plastic and reconstructive surgery*. 1992;90(1):112-118.

50. Kinser A, Sands W, Stone M. Reliability and validity of a pressure algometer. *Journal of Strength and Conditioning Research*. 2009;23(1):312-314.
51. Vaughan B, McLaughlin P, Gosling C. Validity of an electronic pressure algometer. *International Journal of Osteopathic Medicine*. 2007;10(1):24-28.
52. Siemionow M, Zielinski, M., Sari, A. Comparison of clinical evaluation and neurosensory testing in the early diagnosis of superimposed entrapment neuropathy in diabetic patients. *Annals of plastic surgery*. 2006;57(1):9.
53. Tassler PL, Dellon AL. Correlation of measurements of pressure perception using the pressure-specified sensory device with electrodiagnostic testing. *Journal of Occupational & Environmental Medicine*. 1995;37(7):862-866.
54. Uddin Z, MacDermid J, Packham T. The Ten test for sensation. *Journal of Physiotherapy*. 2013;59(2):132-132.
55. Sun HH, Oswald TM, Sachanandani NS, Borschel GH. The Ten test: Application and limitations in assessing sensory function in the paediatric hand. *Journal of Plastic, Reconstructive and Aesthetic Surgery*. 2010;63(11):1849-1852.
56. Faught BE, McKee NH. Establishing a positivity criterion in determining the utility of the Ten test in diagnosing carpal tunnel syndrome. *Journal of Reconstructive Microsurgery*. 2002;18(7):634.
57. Katims JJ, Rouvelas P, Sadler B, Weseley S. Current perception threshold. Reproducibility and comparison with nerve conduction in evaluation of carpal tunnel syndrome. *ASAIO Transactions*. 1989;35(3):280-284.
58. Bleecker ML, Katims JJ, Patil AS, et al. Current perception threshold screening for carpal tunnel syndrome. *Archives of Environmental Health*. Vol 46; 1991.
59. Dotson RM. Clinical neurophysiology laboratory tests to assess the nociceptive system in humans. *Journal of Clinical Neurophysiology*. 1997;14(1):32-45.
60. MacDermid JC, Gross A, Galea V, et al. Developing biologically-based assessment tools for physical therapy management of neck pain. *Journal of Orthopaedic & Sports Physical Therapy*. 2009;39(5):14.
61. Neurometer CPT/C Manual. *Operating Manual* Denver, Colorado: Neurotron, Incorporated; 2010:93.
62. Lowenstein L, Jesse K, Kenton K. Comparison of perception threshold testing and thermal-vibratory testing. *Muscle and Nerve*. 2008;37(4):514-517.
63. Masson E, Veves A, Fernando D, Boulton A. Current perception thresholds: a new, quick, and reproducible method for the assessment of peripheral neuropathy in diabetes mellitus. *Diabetologia*. 1989;32(10):724-728.
64. Park R, Wallace MS, Schulteis G. Relative sensitivity to afentanil and reliability of current perception threshold vs von Frey tactile stimulation and thermal sensory testing. *Journal of the Peripheral Nervous System*. 2001;6(4):232-240.
65. Peters EW, Bienfait HM, de Visser M, de Haan RJ. The reliability of assessment of vibration sense. *Acta Neurologica Scandinavica*. 2003;107(4):293-298.
66. Mahns DA, Perkins NM, Sahai V, Robinson L, Rowe MJ. Vibrotactile frequency discrimination in human hairy skin. *Journal of Neurophysiology*. 2006;95(3):1442-1450.

67. Jensen TS, Bach FW, Kastrup J, Deigaard A, Brennum J. Vibratory and thermal thresholds in diabetics with and without clinical neuropathy. *Acta neurologica Scandinavica*. 1991;84(4):326-333.
68. Chong P, Cros D. Technology literature review: Quantitative sensory testing. *Muscle and Nerve*. 2004;29(5):734-747.
69. van Deursen RW, Sanchez MM, Derr JA, Becker MB, Ulbrecht JS, Cavanagh PR. Vibration perception threshold testing in patients with diabetic neuropathy: ceiling effects and reliability. *Diabetic Medicine*. 2001;18(6):469-475.
70. Mahbub H, Ohnari H, Tanigawa K, et al. Vibrotactile perception at glabrous and nonglabrous skin of fingers: Repeatability of measurements and changes induced by acute vibration exposure. *Journal of Occupational Health*. 2011;53(1):10-15.
71. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A, Allison S. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. *Spine*. 2003;28(1):52-62.
72. Ohrbach R, Gale EN. Pressure pain thresholds, clinical assessment, and differential diagnosis: reliability and validity in patients with myogenic pain. *Pain*. 1989;39(2):157-169.
73. Vatin J-J, Shapira, S., Magora, F. Adler, D. Magora, A. Electronic pressure algometry of deep pain in healthy volunteers. . *Archives of Physical Medicine and Rehabilitation*. 1993;74(5):5.
74. Ylinen J, Nykanen M, Kautiainen H, Hakkinen A. Evaluation of repeatability of pressure algometry on the neck muscles for clinical use. *Manual therapy*. 2007;12(2):192-197.
75. Nussbaum E, Downes, L. Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. *Physical Therapy*. 1998;78(2):9.
76. Jerosch-Herold C. Measuring outcome in median nerve injuries. *The Journal of Hand Surgery: Journal of the British Society for Surgery of the Hand*. 1993;18(5):624-628.
77. Marsh D. The validation of measures of outcome following suture of divided peripheral nerves supplying the hand. *Journal of Hand Surgery (British and European Volume)*. 1990;15(1):25-34.
78. Jerosch-Herold C. Should sensory function after median nerve injury and repair be quantified using two-point discrimination as the critical measure? *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*. 2000;34(4):339-343.
79. Moloney NA, Hall TM, Doody CM. Reliability of thermal quantitative sensory testing: a systematic review. *Journal of Rehabilitation Research & Development*. 2012;49(2):16.
80. Shukla G, Bhatia M, Behari M. Quantitative thermal sensory testing - value of testing for both cold and warm sensation detection in evaluation of small fiber neuropathy. *Clinical Neurology and Neurosurgery*. 2005;107(6):486-490.
81. MacDermid J, Walton D, Cote P, Santaguida L, Gross A, Carlesso L. Use of outcome measures in managing neck pain: An international multidisciplinary survey. *The Open Orthopaedics Journal*. 2013;7(Suppl 4:M10):1-15.

Chapter 2

Test-retest reliability and concurrent validity of the Ten Test

Abstract

Study Design: Cross-sectional, clinical measurement study.

Introduction: The Ten Test (TT) is an easy, fast and quantifiable test of moving light touch sensation requiring no instrumentation, however; evidence regarding its reliability and validity are limited.

Purpose: This study examined the test-retest reliability and concurrent validity of the TT for sensory examination in the hand.

Methods: The Ten Test, current perception threshold (CPT) and vibration perception threshold (VPT) were each measured twice by one rater over one session in 27 volunteers who reported decreased sensation in the hand.

Results: Same day measurements of the Ten Test (TT) demonstrated strong test-retest reliability (ICC: 0.83 – 0.91). Minimal detectable change scores suggest that individual scores need to change by approximately 1 point ($MDC_{90} = 1.57 - 2.15$) to ensure that a score is not the result of measurement error. TT scores did not correlate with CPT or VPT scores.

Conclusions: The TT demonstrated substantial test-retest reliability. These findings support the use of the TT for individual measurements of moving light touch in the research and clinical setting. However, the TT cannot predict CPT and VPT scores.

Level of Evidence: Not applicable.

Test-retest reliability and concurrent validity of the Ten Test

Introduction

Neuromusculoskeletal disorders affect the bones, muscles, joints and nervous system of the body. They are a major burden on Canadians and the health care system. The Institute of Musculoskeletal Health and Arthritis (IMHA) reported that 11 million Canadians were affected by musculoskeletal conditions in 2010, about 3.1% of the population¹. With the aging population, the IMHA anticipates an increase in the prevalence of musculoskeletal conditions to 15 million Canadians in 2031.¹ In the year 2000, the cost of musculoskeletal conditions in Canada was \$22.2 billion, representing the highest total cost of 20 diagnostic categories, followed by cardiovascular diseases and neuropsychiatric conditions.³ Another survey of the general population in Sweden found a prevalence rate of 14.4% (95%CI: 13.0%-15.8%) of numbness, tingling or pain in the median nerve distribution of the hand.⁸²

Sensory threshold tests attempt to determine “how much” sensation a person perceives in an affected area. This is determined by a complicated interaction between the peripheral sensory receptors, peripheral nerves and the central nervous system.¹⁴ Threshold testing typically applies a range of stimuli at standard testing sites to quantify sensory dysfunction.¹

Physical function depends on normal integration between sensory and motor systems.⁴ Sensory evaluation may be employed to diagnose a disorder, identify the severity of sensory impairment, determine the level of axonal regeneration, determine the best course of treatment, establish the need for surgical intervention, identify the progression of a condition and to define the level of hand function.¹⁰ Identifying whether musculoskeletal disorders have a sensory component is critical to accurate diagnosis, treatment, prognosis and outcome evaluation.

There is a dearth of information regarding the usage of sensation testing by clinicians. A recent survey of 381 clinicians from a variety of clinical backgrounds reported that 23% used quantitative sensory testing and 11% used pain algometry routinely in practice as outcome measures for people with neck pain.⁸¹ Another study reported a 67.4% usage rate of sensory tests (specified as monofilament or two-point discrimination tests) by physiotherapists and occupational therapists (n= 242) for use as outcome measures in the population of people with distal radius fracture.⁸³ A third study surveyed hand therapists (n=315) and reported usage rates of 38.2% for the Semmes-Weinstein monofilament test and 17.3% for the two-point discrimination test in people with elbow fractures.⁸⁴

Usage of sensation tests in the clinical setting may be limited due to several factors. First, most sensory tests are psychophysical tests.⁸⁰ They are typically based on subjective responses to quantified stimuli and, hence, results can be contingent on subtle changes in methodology, such as testing site, pressure of stimulator application,

stimulator size and training of subjects.¹¹ The client must be alert, compliant and able to follow instructions.⁶⁸ In fact, responses can only be considered to be valid if the patient is cooperative, and, although testing procedures can be used to show inconsistencies in responses, there is no decisive way to determine whether a client is responding truthfully.¹¹ Second, many of the sensory tests available are limited by unsatisfactory clinical measurement properties, or lack of evidence regarding the same. Further, many sensory measures may take extensive time and, therefore, may be considered burdensome by busy clinicians. Finally, many tests cannot be afforded by clinical practices. Despite the longstanding use of various forms of sensation testing in the clinical and research setting, there is surprisingly limited clinical measurement research on these tools. This limits the ability for clinicians to assess and quantify sensation loss, assess for change in presentation, and provide reliable estimates in assessing sensation.

The Ten Test is a quick and useful clinical test. It requires no special testing equipment, quantifies sensation loss and assesses for change over time. Vibration perception threshold and current perception threshold each take much longer to perform, require trial runs, use expensive equipment and require special training. Albeit, the nerve fibres involved should be similar (large myelinated A- β fibres) for all three tests. Therefore, the Ten Test is a less burdensome test with respect to time and cost. It is a quantitative intensity rating assessment of moving light touch sensation. It provides a ratio of sensation loss compared to a client's own normally innervated comparison site.⁴⁸ To conduct the test, a patient reports the level of sensation from 1 to 10 during light moving touch applied by an examiner's fingertip to the affected area (such as the palmar aspect of the fingertip), while an unaffected, similarly innervated area is simultaneously stimulated.³³ If the symptoms are bilateral, then the lip, cheek or bridge of nose are proposed as alternate reference sites.³³ The Ten Test has promise as an efficient sensory test that would be useful as both an assessment and outcome measure. However, there is limited research into its clinical measurement properties which hinder the confidence in its reliability and validity.

The main purpose of this study was to examine the test-retest reliability and concurrent validity of the Ten Test. The authors expect that the Ten Test will demonstrate strong test-retest reliability and will moderately correlate with current perception threshold (CPT) and vibration perception threshold (VPT). The specific research questions being studied are:

1. What are the same-day test-retest reliabilities, interclass correlation coefficients ($ICC_{2,1}$) and minimal detectable change scores (MDC_{90}) for the Ten Test contrasted with the VPT test and the CPT test in a population of people with decreased sensation in their fingers?

2. What is the concurrent validity (Spearman’s rank correlation coefficient r_s) for the Ten Test, compared to VPT and the CPT tests in a population of people with decreased sensation in their fingers?

Methods

Subjects

E Twenty-seven subjects were recruited (see Table 1 for participant demographics). This approval was obtained from the Hamilton Integrated Research Ethics Board. All participants reported experiencing decreased sensation in one or both hands. Inclusion criteria consisted of subjects over 16 years of age, decreased sensation in the hands/fingers, able to provide consent, and able to participate in English verbally and in writing. Exclusion criteria included participating in a drug trial within the last 3 months. Participants were recruited by public advertisement and also were recruited from previous participation in research trials at the MacHand Lab, McMaster University. Participants represented a wide range of diagnoses such as carpal tunnel syndrome, fibromyalgia, post-chemotherapy neuropathy and ulnar neuropathy, and a wide range of ages.

Table 1. Participant Characteristics (n = 27, SD= standard deviation)

Characteristics			
Age (years)	Mean	SD	Range
	53.0	22.4	56 (min 19, max 86)
Sex (%)	Male	Females	
	7 (25.9%)	20 (74.1%)	
Symptoms (%)	Unilateral	Bilateral	
	10 (37%)	17 (63%)	
Most affected (%)	Left	Right	
	18 (67%)	9 (33%)	

Key: n = sample size; SD = standard deviation

Procedure

The study procedure involved a single one hour session for each participant. Participants were pre-screened for inclusion and exclusion criteria via a phone call or email prior to scheduling an appointment. The purpose of the study and the potential risks were explained to all participants and consent and demographic forms were completed prior to implementing the assessments.

Standardized examination of the Ten Test, VPT and CPT testing was performed. All participants were assessed by a single rater (ND), who is a physiotherapist and the principal investigator of the study. The rater received training on the three testing

methods, including practice sessions to ensure competence. The testing order was: the Ten Test, followed by a randomized order of CPT and VPT, followed by repeated testing of the CPT and VPT in the randomized order, and then repeated testing of the Ten Test. The Ten Test was administered as the first and last test to minimize the potential for participants to recall their initial responses. Testing was performed on the affected hand, or, in the case of bilateral symptoms, the most affected hand. For CPT and VPT, tests were performed on D1 (thumb), D2 (index finger) and D5 (little finger). For the Ten Test, all fingers were assessed, as per protocol, but values were recorded solely for D1, D2 and D5. For all tests, subjects were blinded to test results. It was not possible to blind the rater.

The Ten Test.

Each participant was assessed using the Ten Test according to established protocols based on those proposed by the originator of the test.³³ Participants were seated comfortably at a table and their forearms and hands were supported on a pillow in a supinated position to easily expose the palms. Moving light touch was first applied by the examiner's fingertip to an area perceived as normal. In participants with unilateral symptoms, the contralateral fingertip was used as the comparison site. When bilateral symptoms were reported, another comparison site was selected, often a hairless area of the distal forearm. The 'normal' area was defined as a 10 on an analogue scale of 1 to 10. A fingertip on the affected hand was then simultaneously stroked with equal pressure and the participants were asked to rate their sensibility in relation to the unaffected area, where 1 equals no sensation and 10 is normal sensation. Responses were recorded. As per the protocol², the test was applied to each fingertip on the palmar aspect and subjects were not blinded.

Current perception threshold.

CPT was assessed with the Neurometer[®] (Neurotron, Inc, Baltimore, Maryland, http://www.neurotron.com/Neurometer_CPT-C.html). CPT assesses the function of sensory nerves with the use of low levels of electrical current.⁵⁸ A microprocessor-controlled electrical stimulator delivers a constant alternating current at a choice of three different frequencies: 5Hz, 250Hz or 2000Hz.⁵⁹ The constant current delivery compensates for variations in skin resistance, skin thickness and moisture.⁵⁸ The Neurometer[®] provides a minimal stimulus using an automated protocol that can be administered following little training.⁶⁰ The three different frequencies assess three types of afferent neurons: large myelinated A- β fibres (2000Hz), A- δ fibres (250Hz) and C fibres (5 Hz).⁶² Comparisons with nerve conduction studies, thermal sensory testing, vibration perception threshold, somatosensory evoked potential studies, histopathological and pharmacological studies support the concept of nerve fibre selectivity based on frequency.⁵⁹ Normal range values are 6 - 13 (no units), values less

than 6 are considered to represent hypersensitivity, and values higher than 13 indicate a loss of sensation, or hyposensitivity.⁶¹

The participants were seated comfortably at a table. One centimeter diameter, gold-plated, gel-coated electrodes were applied to the medial and lateral distal fingertip of the tested digit and secured with adhesive tape. The procedure was explained to the participant and a trial application was performed to allow participants to familiarize themselves with the test. Participants used a control device to initiate the onset of delivery of stimulation and to respond at the moment that sensation was perceived. Participants were unable to read the intensity of current applied during testing. Stimulations of 2000Hz were applied; as this frequency has been shown to stimulate nerve fibres associated with moving light touch.⁵⁹ Assessments were performed for the thumb, index and little finger to isolate the median and ulnar nerve cutaneous distributions. The CPT for each finger was determined by using the range-CPT protocol. The mean value after receiving three consistent responses is reported.

Vibration perception threshold.

For the assessment of VPT, we used the TSA-II NeuroSensory Analyzer (Medoc Ltd., Israel, <http://www.medoc-web.com/products/tsa-ii>). Participants were seated comfortably at a table with their affected forearm supinated and hand resting on a pillow, in a position where the computer screen could not be seen. Although VPT assessment in the hand is commonly assessed at the distal palmar fingertip surface, this area is prone to callouses, and it is not known to what degree this affects measurement.¹⁶ Age has been disputed as having an effect on VPT thresholds in the hand, and may need to be accounted for in interpreting results, although height and gender consistently have not impacted VPT in the hand, as it does for the foot.^{16, 65} The delivery of vibration is standardized, although parameters can be altered. Our trial used a 100Hz frequency. We slowed the ramp speed to decrease the error potential created by response time delay. Differences occur in the vibration perception thresholds of glabrous (non-hairy) and non-glabrous (hairy) skin, with greater discriminative power of glabrous skin.⁶⁶ Therefore, testing should be consistent in assessing VPT at glabrous skin sites. The method of limits was used, where an increasing intensity of vibration was delivered through an applicator head with a 1.22cm² surface area until the participant sensed vibration, at which time a button on a computer mouse was pressed to indicate the perception threshold (in microns). A practice trial was performed, as decreased reliability in the first trial has been shown.¹² Assessments were performed at the tip of the thumb, index and little finger, to minimize the risk of overlapping nerve distributions. Eight trials were performed at each test site, the standard protocol for this device. The mean of the trials was recorded as the VPT (in microns). Normal range values for this device are not available.

Statistical Analysis

The target sample size was based on a graphical sample size calculator that indicated that high reliability can be detected with a sample size of 40 participants.⁸⁵ Unfortunately, only 27 subjects enrolled. The description of participants is described using mean (standard deviation) for continuous variables and counts for categorical variables (see Table 1).

Test-Retest Reliability - Descriptive statistics including the means and standard deviations for each trial were calculated for the Ten Test, CPT and VPT tests for the total sample (n=27). Reliability analysis was performed to estimate the test-retest reliability of the Ten Test, VPT and CPT when used in the hand. Interclass correlation coefficients for single measures (ICC_{2,1}) for the Ten Test, VPT and CPT, with the 95% confidence intervals (95%CI), and standard error of measurement (SEM) were calculated. There is no standard agreement on acceptable levels of ICC values for test-retest reliability.⁸⁶ It has been suggested that ICC values for test-retest reliability have a minimum score of 0.60 to be clinically useful;⁸⁷ however, it is more useful to judge the value of a test based on the comparison of reliability scores of alternative tests. Minimal detectable change at the 90% confidence level (MDC₉₀) was calculated as follows: SEM x 1.645 $\sqrt{2}$. Scores for each participant were calculated for the first, second and fifth digits. Bland and Altman plots were created. They present a scatter plot showing the difference between test 1 and test 2 against their means. The 95% confidence bands were also drawn (mean difference \pm 1.96 SD_{diff}). Bland and Altman plots were also evaluated for systematic bias and heteroscedasticity. Systematic bias will be observed if the scatter plot demonstrates the mean differences being largely negative or positive. Further, if zero is included in the 95% CI bands, then no significant change of the mean has occurred. Heteroscedasticity is detected when larger variability exists for either lower or higher test values.

Concurrent Validity – The data from the first trial of each test was used for analysis. Evaluations were made for the first, second and fifth digits. Linear regression was used to describe the relationship between each pair of test scores. Data was first plotted, and a Spearman's rank correlation coefficient (r_s) was calculated. Minimal significance level was set at $p < 0.05$. Strengths of Spearman's correlation have been defined as ≤ 0.35 weak, 0.36 to 0.67 moderate, and 0.68 to 1.0 strong (Weber & Lamb, 1970; Mason, Lind & Marchal, 1983; as cited in Taylor, 1990).⁸⁸ All analyses were performed using SPSS 20 for Windows (SPSS Inc., Chicago, IL).

Results

Study participants had a mean (\pm SD) age of 53 (\pm 22) years (see Table 1). Approximately one quarter of participants was male and 3/4 female. Almost two thirds (63%) of participants had bilateral symptoms. The left hand was primarily affected in just over 2/3 of cases. A variety of diagnoses were represented, such as ulnar neuropathy, carpal tunnel syndrome, spinal stenosis, fibromyalgia and following severe frostbite.

It can be observed that most subjects perceived mild sensory loss (mean Ten Test score 7.93-8.20/10) (see Table 2). Most of the scores demonstrated a high level of skewness (coefficients of skewness less than -1 or greater than +1). Further, most coefficients of skewness for the Ten Test and VPT had symmetries which differed from that of a normal distribution. The coefficients of kurtosis demonstrated that, with the exception of the D2 and D5 VPT values, the peakedness of the distribution curves were not significantly different than that of a normal distribution.

Table 2. Mean scores with standard deviations (SD)

Test	Median test 1	Mean test 1	SD	Median test 2	Mean test 2	SD	Coefficient of skewness	Coefficient of kurtosis
D1								
TT	9	7.93	2.77	9	7.48	2.97	-1.20*	0.29
CPT	14	13.92	5.93	14	13.65	5.91	-0.40	-0.07
VPT	0.76	0.96	0.82	0.62	1.07	0.99	1.57*	1.37
D2								
TT	9	8.04	2.14	9	7.74	2.40	-1.02*	-0.19
CPT	12	11.48	5.42	11	10.88	5.28	-0.58	-0.27
VPT	0.55	1.06	1.12	0.60	1.11	1.27	1.73*	2.12*
D5								
TT	9	8.20	2.24	9	7.93	2.32	-1.21*	0.74
CPT	10	10.70	4.31	10	10.93	5.25	0.13	0.05
VPT	1.01	1.40	1.34	0.77	1.38	1.30	2.04*	4.91*

*indicates that distribution is significantly different from a normal curve

TT = Ten Test

CPT = current perception threshold

VPT = vibration perception threshold

SD = standard deviation

D1 = thumb

D2 = index finger

D3 = little finger

Test-Retest Reliability. Table 3 shows the means (SD) for test 1 and test 2, single measure $ICC_{(2,1)}$ values with 95% CI, level of significance, SEM and MDC_{90} values for D1, D2 and D5 for the Ten Test, CPT and VPT. Ten Test and VPT scores had better repeatability than CPT scores.

Ten Test test-retest reliability: The $ICC_{(2,1)}$ values were strong overall for the TT (ICC: 0.83 – 0.91). Comparing the three fingers, D5 had lower ICC scores for both the Ten Test and CPT. The MDC_{90} scores show that, for the TT, individual scores need to change by approximately 2 points (min 1.57 – max 2.15) to ensure that a patient’s score is not the result of measurement error.

CPT test-retest reliability: CPT reliability scores were moderate to strong (ICC: 0.63 -0.83) and had large confidence intervals.

VPT test-retest reliability: The ICC values found in this study (ICC: 0.88 to 0.95) for VPT were high.

Table 3. Test-retest reliability

Test	ICC	95%CI	SEM	MDC_{90}
D1				
TT	0.91	0.82-0.96	0.83	1.93
CPT	0.77	0.54-0.89	2.84	6.61
VPT	0.89	0.77-0.95	0.27	0.63
D2				
TT	0.90	0.80-0.96	0.68	1.57
CPT	0.83	0.66-0.92	2.23	5.20
VPT	0.88	0.75-0.95	0.38	0.90
D5				
TT	0.83	0.66-0.92	0.92	2.15
CPT	0.63	0.34-0.81	2.62	6.10
VPT	0.95	0.88-0.98	0.30	0.70

TT = Ten Test

CPT = current perception threshold

VPT = vibration perception threshold

SD = standard deviation

95%CI = 95% confidence interval

MDC = minimal detectable change

D1 = thumb

D2 = index finger

D3 = little finger

SEM = standard error of the mean

$ICC_{(2,1)}$ = Single measure interclass correlation coefficient

Bland and Altman plots: From the scatter plots, it can be observed that no systematic bias was observed for any of the tests, as the differences between trials were not predominately positive or negative, and the 95% confidence bands included zero (Figures 1-3 are provided for D1 of each test). Heteroscedasticity was not observed for Ten Test or CPT scores. Heteroscedasticity was observed for VPT, where a greater spread of data was observed at the higher end of the scales (which represents greater sensory abnormality).

Figure 1. Bland and Altman plot. *Note:* several subjects had identical values.

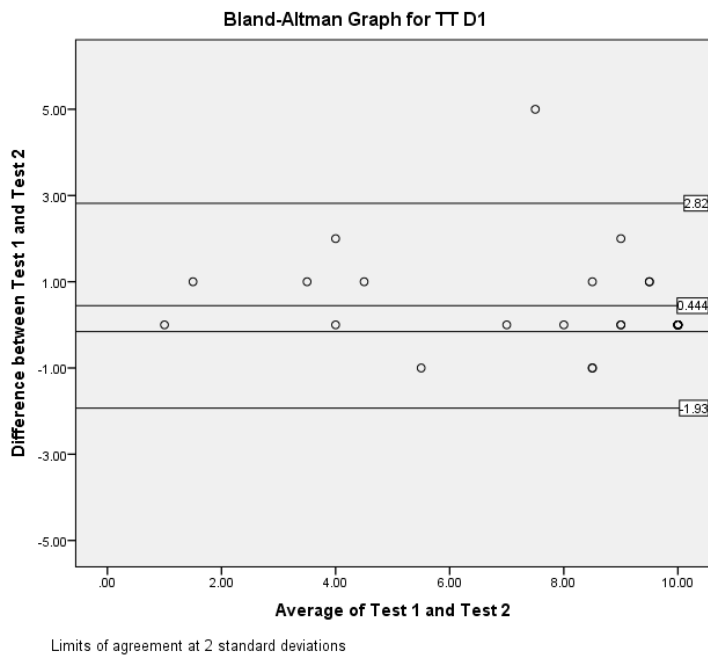


Figure 2. Bland and Altman plot.

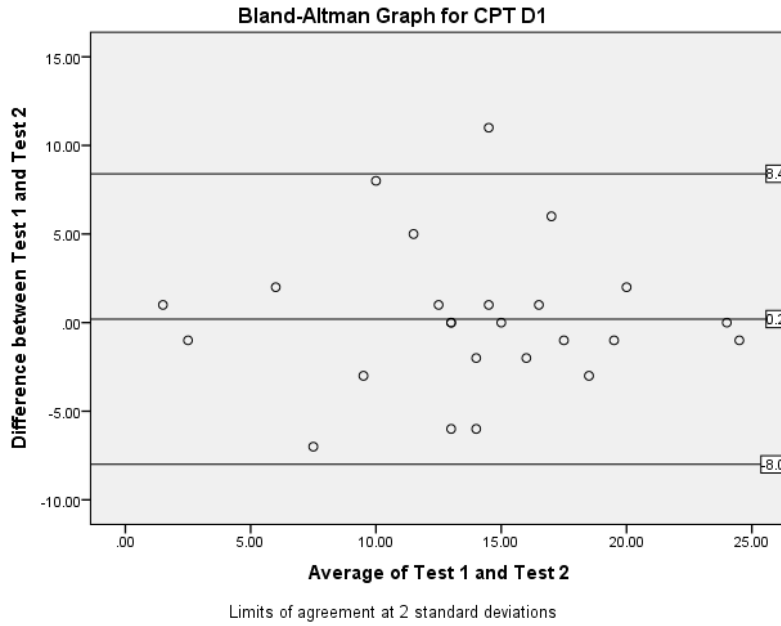
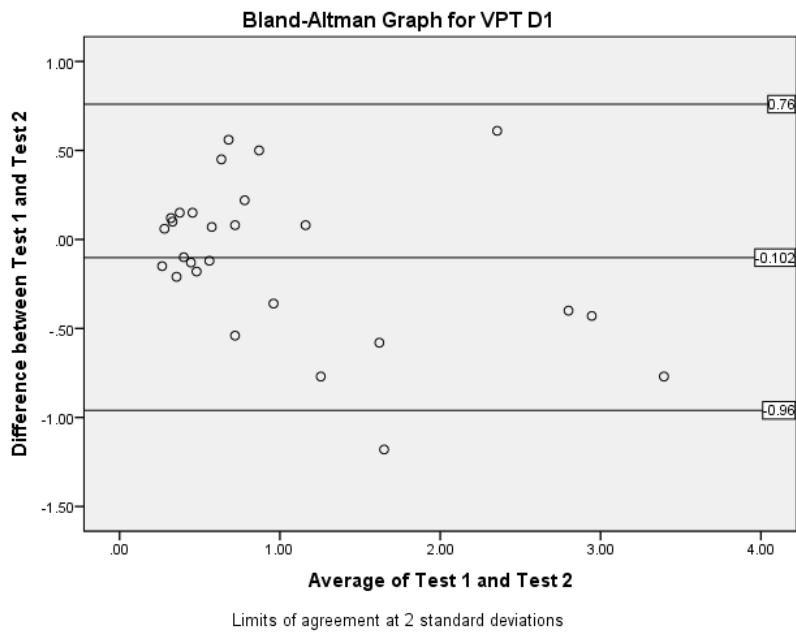


Figure 3. Bland and Altman plot.



Concurrent Validity. There was no statistically significant correlation between the Ten Test and VPT or CPT scores for any of the fingers tested ($r_s = -0.30$ to 0.16 , $n = 27$) (see Table 4). If the Ten Test was correlated with CPT or VPT, negative values would be expected, as greater sensory loss is represented by lower Ten Test scores and higher VPT and CPT scores. Both positive and negative values were observed. The only statistically significant correlation was observed between CPT and VPT scores for D5 (Spearman’s $r_s = 0.40$, $p = 0.04$).

Table 4. Correlations between the Ten Test, current perception threshold and vibration perception threshold.

Test	D1			D2			D5		
	TT	CPT	VPT	TT	CPT	VPT	TT	CPT	VPT
TT	-	-0.02 (0.92)	-0.30 (0.13)	-	0.01 (0.97)	-0.26 (0.20)	-	0.00 (0.99)	0.16 (0.43)
CPT	-	-	0.24 (0.24)	-	-	0.15 (0.46)	-	-	0.40* (0.04)
VPT	-	-	-	-	-	-	-	-	-

Key: Note: Correlation values represent Spearman’s rank correlation coefficient r_s values with p values presented in brackets.

*represents a statistically significant correlation

TT = Ten Test

D1 = thumb

CPT = current perception threshold

D2 = index finger

VPT = vibration perception threshold

D3 = little finger

The Ten Test has strong test-retest reliability in adults with perceived sensory loss and acceptable MDC_{90} scores; however, it may not identify sensory changes in asymptomatic individuals. Ten Test scores did not correlate with CPT or VPT scores.

Discussion:

Clinicians should use the Ten Test in their practice for people with sensory deficits (see Figure 4). We found the Ten Test to be responsive to change; the CPT test-retest reliability to be moderate and the VPT test-retest reliability to be high. There was no correlation between the Ten Test with CPT or VPT.

Figure 4: The Ten Test



The Ten Test was responsive to change. In comparison, the MDC for the widely used numerical pain rating scale is 4.1 points on an eleven point scale (0-10) in the population of people with cervical radiculopathy³ and 1.1 points in the population of people with shoulder pain.⁴

No research on test-retest reliability of the Ten Test was identified. Previous research has demonstrated high inter-rater and low to high intra-rater reliability.² Strauch et al.² tested for inter-rater reliability between two raters, and found an interclass correlation coefficient of ICC = 0.91 (95%CI: 0.87 to 0.94, $p < 0.05$), demonstrating excellent reliability between examiners.

Inter-rater reliability in the pediatric population was found to have a kappa agreement of 1.0⁵. Strauch et al.² tested for inter-rater reliability between two raters in an adult population with peripheral nerve disorders, and found an interclass correlation coefficient of 0.91 (95% CI 0.87 to 0.94, $p < 0.05$), demonstrating excellent reliability between examiners. The intra-rater reliability of six raters was good for four out of six raters (ICC 0.61 to 0.90), however; two raters had ICC values of 0.25 and 0.38.²

Several limitations in the assessment of test-retest reliability are noted. As the repeated testing occurred on the same day, there was a potential for recall bias. The authors attempted to reduce this by assessing the Ten Test first and last in the testing sequence, in order to increase the time and distraction between repeated Ten Test trials. In pre-data collection simulations, volunteer participants reported that they were unable to recollect their initial responses using this method. However, if these two separate evaluations were not sufficiently separated in time to avoid recall, the “test retest” reliability might be considered to indicate intra-rater reliability, rather than separate occasions. Typically, intra-rater reliability is higher than test retest reliability.

In our study, there was a large set of data collected for “normal” fingers – that is, without perceived loss of sensation. Since D1, D2 and D5 were tested and scores recorded for all fingers, many scores were recorded that would have represented

normally innervated fingers. For example, if a participant had carpal tunnel syndrome, D5 scores would likely represent normal findings.

The main limitation in this study is the small sample size of 27 subjects. This will contribute to a lack of precision in our estimates. Wide confidence intervals were present in a number of our reliability estimates, with the exception being estimates where the ICC point estimate was very high. Thus, our confidence in the reliability estimate is better for the TT than for the other measures. Since convenience sampling was used, a lack of volunteerism is the primary reason for the low sample size. Our testing protocol was relatively low burden and we used multiple strategies for recruitment. The lack of response may indicate that sensory deficits are harder for potential subjects to identify than pain or mobility deficits.

The CPT test-retest was moderately reliable with large confidence intervals. Previous studies have reported statistically significant differences in test scores between repeated testing.^{6,7} Two studies found coefficients of variation ranging between 6.5 - 11% for CPT testing at 2000Hz on repeated testing.^{8,9} This study is consistent with previously studies in identifying concerns regarding test-retest reliability for CPT.

The VPT test-retest reliability was high. These results are very similar to previous studies, which also found high VPT test-retest reliability in populations of healthy adults, people with diabetes and peripheral neuropathy (foot), and spinal cord injury.^{6, 10-12} This study adds to the population of people with hand sensory deficits to populations studied for VPT reliability.

We found no convincing evidence of concurrent validity of the Ten Test with CPT and VPT. This suggests that the tests are measuring different constructs, and that the scores from one test cannot be used to predict results from another. This was unexpected, as all tests are expected to activate similar nerve fibres (large myelinated A- β fibres).^{6, 13} In contrast, Strauch et al.² tested for concurrent validity between the TT and the Semmes-Weinstein monofilament test (SWMT). Results produced Spearman's rank correlation coefficient of $r_s = -0.71$ (95% CI -0.68 to -1.0, $p < 0.05$). This higher association may be partly explained by the Ten Test and SWMT being two ordinal scales (with 10 and 5 points, respectively), thereby more easily producing a correlation effect. In this study, the comparison was performed between the Ten Test and a 25 point ordinal scale (CPT), and a ratio scale (VPT). Furthermore, as our subjects tended to have mild sensory deficits, as seen by mean Ten Test scores of 9/10, and mean CPT scores of 14 (decreased sensation) and 11 (normal sensation), they represented a narrow range of the spectrum of sensory disturbance. Examining correlations over a narrow range may not produce the same extent of correlation as when a wider spectrum is examined.

Another contributing factor for the lack of association between test scores may be related to testing methods. CPT stimulates the nerve fibres directly as the current intensities delivered are below that which would stimulate skin receptors.¹⁴ VPT activates sensory receptors. The Ten Test reflects a person's perception of sensation and is dependent on what is perceived as "normal" and their ability to quantify a stimulus. Thus, the lack of correlation between the tests may be a result of the tests primarily evaluating different components of the sensory pathways. As well, while CPT and VPT are threshold tests, the Ten Test is an intensity rating scale, and therefore is subject to a different cognitive evaluation process. The cognitive process in evaluating sensation as being normal is potentially more foreign or challenging to subjects compared to rating pain on a numeric scale. Numeric pain rating scales have been shown to be highly reliable,^{3,4} so the construct of numeric rating for sensation is viable, but potentially more challenging when evaluating sensory perception. Further, numeric pain rating scales often use a rating of 0 to indicate no pain, which may be more intuitive for subjects than a rating of 1.

A further possible reason for the lack of correlation between test scores is the moderate test-retest reliability observed for the CPT scores. Additionally, light moving touch, current perception and vibration are clearly not the same sensations. That may have been a factor in the lack of correlation found. However, research has shown that there are additional sensory signals via application force, application technique, and vibration through the examiners hand that recruit other types of sensory receptors and, therefore, test selectivity cannot be assumed.¹⁵

In comparison to this study's results, previous research has demonstrated moderate correlation between CPT at 2000Hz and VPT (Spearman's rank correlation coefficient $r_s = 0.42 - 0.50$).^{6,8} Our finding of moderate correlation ($r_s = 0.40$, $p = 0.04$) for D5 is similar. Our findings of no correlation for D1 and D2 may be partly due to the small sample size and different population studied, with the previous study populations of, 1) healthy adults and, 2) adults with diabetes.

Limitations of the Ten Test. Several limitations of the Ten Test were noted. Although not evident in the data presented, it was likely that the Ten Test sometimes failed to identify decreased sensibility. For example, several participants who reported 10/10 on the Ten Test received very high VPT and CPT scores. One example is a 67 year old male. He reported 10/10 for the TT on all fingers, yet had VPT scores of 0.61 (D1), 1.35 (D2) and 6.08 (D5). The values for D2 and D5 are much higher than the median values observed in this sample (D2 = 0.55; D5 = 1.01). This may be explained by the reliance on perception which is inherent in the Ten Test. For example, if a person has had gradual and bilateral loss of sensation, they may have become accustomed to the loss, and thus perceive the reduced sensory capacity as normal. The Ten Test may have a high false negative rate (type 2 error), an important clinical consideration.

In contrast, one previous research paper reported that the Ten Test was able to detect minimal loss of sensation in patients with carpal tunnel syndrome, superior to the WEST, static, and moving two point discrimination tests¹⁶. However, the SWMT was used the gold standard criterion, and the ability of the SWMT to identify asymptomatic loss of sensibility has not been established. Further, two-point discrimination tests have been shown to be poor at identifying early loss of sensation^{17, 18}. This study used VPT. It is a superior criterion in evaluating loss of sensibility, and has been shown to be more sensitive than the SWMT (Sorman and Edwall, 2002, as cited in Wu, Driver, Wrobel and Armstrong, 2007).¹⁹ In fact, increased VPT is one the first signs of peripheral neuropathies, nerve entrapment and reinnervation following repair.^{20, 21}

Another limitation of the Ten Test is that it may not isolate true innervation points, as the area covered in the moving light touch may cross over more than one sensory field. Also, the protocol for the Ten Test specifies that each of the ten fingertips are tested, but omits the consideration of a test point that would isolate the radial nerve, such as just proximal to the dorsal aspect of the 1st metacarpophalangeal joint. However, a clinician could easily modify the Ten Test to use for any isolated nerve distribution required.

A further limitation is that the Ten Test protocol does not allow for assessing hyperesthesia, although hyperesthesia has been documented by some authors as a value $>10/10$ ¹⁶ or altering the reference, using the 1 to 10 scale, where 1 = normal, and higher numbers represent increasing hyperesthesia²², although it is unclear what the reference value of 10 would represent on this scale and may be too ambiguous.

Much of the reliability and validity data has been reported by the developer and additional independent testing is needed. As a subjective test, rating the perception of sensibility may be challenging for some people.²²

Conclusion:

The TT is a reliable and clinically useful test for decreased sensation in the hand in people with sensory impairments. It cannot be expected to predict VPT or CPT scores. VPT is a more sensitive tool to early sensation loss and also has strong test-retest reliability scores; however, it is not likely to be readily available in the clinical setting. Further research of test-retest reliability for the Ten Test, especially with larger populations and tested over a multi-day interval, would improve upon the estimate of test-retest reliability. Further assessment of validity is needed to add to the sparse literature for this test.

References

1. The Institute of Musculoskeletal Health and Arthritis. Facts & figures. May 14, 2013; <http://www.cihr-irsc.gc.ca/e/27028.html>. Accessed August 8, 2013.
2. Public Health Agency of Canada. Cardiovascular disease - Economic burden of illness. April 22, 2013; http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/cvd_ebic-mcv_femc-eng.php. Accessed August 8, 2013.
3. Atroshi I, Gummesson C, Johnsson R, Ornstein E, Ranstam J, Rosen I. Prevalence of carpal tunnel syndrome in a general population. *The Journal of the American Medical Association*. 1999;282(2):153-158.
4. Bear-Lehman J, Abreau B. Evaluating the hand: issues in reliability and validity. *Physical Therapy*. 1989;69(12):11.
5. Warwick D, Dunn R, Melikyan E, Vadher J. Assessment. *Hand Surgery: Oxford Specialist Handbooks in Surgery*; 2009:1-37.
6. Jerosch-Herold C. A study of the relative responsiveness of five sensibility tests for assessment of recovery after median nerve injury and repair. *The Journal of Hand Surgery: British & European Volume*. 6 2003;28(3):255-260.
7. Jerosch-Herold C. Assessment of sensibility after nerve injury and repair: a systematic review of evidence for validity, reliability and responsiveness of tests. *The Journal of Hand Surgery: British & European Volume*. 6 2005;30(3):252-264.
8. MacDermid J, Walton D, Cote P, Santaguida L, Gross A, Carlesso L. Use of outcome measures in managing neck pain: An international multidisciplinary survey. *The Open Orthopaedics Journal*. 2013;7(Suppl 4:M10):1-15.
9. Michlovitz SL. Distal radius fractures: Therapy practice patterns. *Journal of Hand Therapy*. 2001;14(4):8.
10. MacDermid JC, Vincent JI, Kieffer L, Kieffer A, Demaiter J, MacIntosh S. A survey of practice patterns for rehabilitation post elbow fracture. *The Open Orthopaedics Journal*. 2012;6:11.
11. Shukla G, Bhatia M, Behari M. Quantitative thermal sensory testing - value of testing for both cold and warm sensation detection in evaluation of small fiber neuropathy. *Clinical neurology and neurosurgery*. 2005;107(6):486-490.
12. Zaslansky R, Yarnitsky D. Clinical applications of quantitative sensory testing (QST). *Journal of the neurological sciences*. 1998;153(2):215-238.
13. Chong P, Cros D. Technology literature review: Quantitative sensory testing. *Muscle & nerve*. 2004;29(5):734-747.
14. Strauch B, Lang A, Ferder M, Keyes-Ford M, Freeman K, Newstein D. The Ten Test. *Plastic and reconstructive surgery*. 1997;99(4):1074-1078.
15. Strauch B, Lang A. The Ten Test revisited. *Plastic & Reconstructive Surgery*. 2003;112(2):593-594.
16. Bleecker ML, Katims JJ, Patil AS, et al. Current perception threshold screening for carpal tunnel syndrome. *Archives of Environmental Health*. Vol 46; 1991:207+.

17. Dotson RM. Clinical neurophysiology laboratory tests to assess the nociceptive system in humans. *Journal of clinical neurophysiology*. 1997;14(1):32-45.
18. MacDermid JC, Gross A, Galea V, et al. Developing biologically-based assessment tools for physical therapy management of neck pain. *Journal of Orthopaedic & Sports Physical Therapy*. 2009;39(5):14.
19. Lowenstein L, Jesse K, Kenton K. Comparison of perception threshold testing and thermal-vibratory testing. *Muscle & nerve*. April 2008 2008;37(4):514-517.
20. Neurometer CPT/C Manual. *Operating Manual* Denver, Colorado: Neurotron, Incorporated; 2010:93.
21. Gerr F, Letz R. Vibrotactile threshold testing in occupational health: A review of current issues and limitations. *Environmental research*. 1993;60(1):145-159.
22. Peters EW, Bienfait HM, de Visser M, de Haan RJ. The reliability of assessment of vibration sense. *Acta Neurologica Scandinavica*. 2003;107(4):293-298.
23. Mahns DA, Perkins NM, Sahai V, Robinson L, Rowe MJ. Vibrotactile frequency discrimination in human hairy skin. *Journal of Neurophysiology*. 2006;95(3):1442-1450.
24. Grunert BK, Wertsch JJ, Matloub HS, McCallum-Burke S. Reliability of sensory threshold measurement using a digital vibrogram. *Journal of Occupational Medicine*. 1990;32(2):100-102.
25. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Statistics in medicine*. 1987;6(4):441-448.
26. Bruton A, Conway J, Holgate S. Reliability: What is it, and how is it measured? *Physiotherapy*. 2000;86(2):6.
27. Chinn S. Repeatability and method comparison. *Thorax*. 1991;46(6):3.
28. Taylor R. Interpretation of the correlation coefficient: A basic review. *Journal of diagnostic medical sonography*. 1990;6(1):5.
29. Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the Neck Disability Index, Patient-Specific Functional Scale, and Numeric Pain Rating Scale in patients with cervical radiculopathy. *American Journal of Physical Medicine & Rehabilitation*. 2010;89(10):9.
30. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. *Journal of Shoulder and Elbow Surgery*. 200911/12 2009;18(6):920-926.
31. Sun HH, Oswald TM, Sachanandani NS, Borschel GH. The Ten Test: Application and limitations in assessing sensory function in the paediatric hand. *Journal of Plastic, Reconstructive & Aesthetic Surgery*. 201011 2010;63(11):1849-1852.
32. Park R, Wallace MS, Schulteis G. Relative sensitivity to alfentanil and reliability of current perception threshold vs von Frey tactile stimulation and thermal sensory testing. *Journal of the Peripheral Nervous System*. 2001;6(4):232-240.
33. Masson E, Veves A, Fernando D, Boulton A. Current perception thresholds: a new, quick, and reproducible method for the assessment of peripheral neuropathy in diabetes mellitus. *Diabetologia*. 1989;32(10):724-728.

34. Katims JJ, Rouvelas P, Sadler B, Weseley S. Current perception threshold. Reproducibility and comparison with nerve conduction in evaluation of carpal tunnel syndrome. *ASAIO transactions*. 1989;35(3):280-284.
35. van Deursen RW, Sanchez MM, Derr JA, Becker MB, Ulbrecht JS, Cavanagh PR. Vibration perception threshold testing in patients with diabetic neuropathy: ceiling effects and reliability. *Diabetic Medicine*. 2001;18(6):469-475.
36. Mahbub H, Ohnari H, Tanigawa K, et al. Vibrotactile perception at glabrous and nonglabrous skin of fingers: Repeatability of measurements and changes induced by acute vibration exposure. *Journal of Occupational Health*. 2011;53(1):10-15.
37. Felix ER, Widerstrom-Noga EG. Reliability and validity of quantitative sensory testing in persons with spinal cord injury and neuropathic pain. *Journal of Rehabilitation Research and Development*. 2009;46(1):69-83.
38. Yarnitsky D. Quantitative sensory testing. *Muscle & Nerve*. 1997;20(2):198-204.
39. Bell-Krotoski JA, Buford WL. The force/time relationship of clinically used sensory testing instruments. *Journal of Hand Therapy*. 1997;10(4):297-309.
40. Patel MR, Bassini L. A comparison of five tests for determining hand sensibility. *Journal of reconstructive microsurgery*. 1999;15(7):4.
41. Boscheinen-Morrin J, Conolly B. Assessment. *The Hand: Fundamentals of Therapy*. Vol 3rd: Butterworth-Heinemann; 2001:1-12.
42. Wu SC, Driver VR, Wrobel JS, Armstrong DG. Foot ulcers in the diabetic patient, prevention and treatment. *Vascular Health Risk Management*. 2007;3(1):12.
43. Dellon A. *Evaluation of sensibility and re-education of sensation in the hand*. Baltimore, Md: Williams & Wilkins; 1981.
44. Uddin Z, MacDermid J, Packham T. The Ten Test for sensation. *Journal of Physiotherapy*. 2013;59(2):132-132.

The validity and patient preference of a visual version of the Ten Test

Abstract

Study Design: Cross-sectional validity and patient preference survey

Objectives: To examine the validity of a new, visual version of the Ten Test (TT) and explore patient preferences between the verbal and visual versions in people with decreased hand sensation.

Background: Hand sensation deficits are common; however, clinically useful quantitative tools which assess sensation are limited. The Ten Test has been found to be a reliable and valid test of sensation in the hand. A visual version was created and concurrent validity was assessed, as well as patient preferences between the two versions.

Methods: Participants (n=14) with impaired sensation in the hand completed the Ten Test, current perception threshold, vibration perception threshold and a visual version of the Ten Test over one session. Spearman's rank correlation coefficient (r_s) was calculated for D(digit)1, D2 and D5. Minimal significance level was set at $p = 0.05$. A percentage was calculated to reflect participant preference between the original and visual TT versions.

Results: The scores for the visual version of the TT correlated strongly with the original version of the TT (Spearman's rank correlation coefficient $r_s = 0.74 - 0.90$). There was no correlation between the visual or verbal versions of the TT with VPT and CPT scores. Participants strongly preferred the visual version of the TT (85.7%).

Conclusions: The visual version of the TT is a valid substitute for the classic verbally administered version of the TT and is preferred by patients.

Key Words: TT, validity, sensation threshold testing

The validity and patient preference of a visual version of the Ten Test

Introduction

Sensory abnormalities in the hand are common. For example, a 1997 survey of 2466 people representing the general population in Sweden reported incidence rates of pain and numbness or tingling of the median nerve distribution of the hand at 14.4% (95%CI: 13.0, 15.8%).¹ The costs associated with musculoskeletal conditions are due to direct treatment costs such as hospital, medication and physician costs, as well as indirect costs such as mortality, and long- and short-term disability.¹

Accurate tools for sensory evaluation are needed in order for clinicians to diagnose disorders of the nervous system, identify severity of impairment, determine the level of regeneration, determine the best course of treatment, delineate the need for surgical intervention, identify the disorder progression, establish the level of hand function, evaluate treatment effectiveness and develop a related treatment plan.^{2,3} Outcome measures decrease the subjectivity of clinical assessment.⁴ Clinicians and researchers require sensation tests that have good reliability and validity. Ideally, measures are able to assess health status change over time, clinically meaningful change, be sensitive in assessing a wide range of deficit severity and be standardized.⁵ Although sensory impairment is common, the utilization of sensory tests may be limited by burden of time, limited information on clinical measurement properties, equipment requirements, and skill required in administering them.

Strauch et al. described the Ten Test (TT) in 1997⁶ as an efficient sensory evaluation tool and outcome measure requiring no instrumentation. It is a verbally administered numeric intensity rating measurement of moving light touch sensation. It produces a ratio of scores derived by comparing sensation in an affected area to the subject's own normally innervated comparison site.⁷

There are few previous studies which have examined the clinical measurement properties of the TT; however, they have produced promising results. The TT has been examined for concurrent validity to the Semmes-Weinstein monofilament test, reporting findings of a Spearman's rank correlation coefficient of $r_s = -0.71$ (95%CI: 0.68, 1.0, $p < 0.05$). The TT has been reported to be functional in the pediatric population over five years of age (for example, at age 5 years old, >80% of children were able to successfully complete the TT).⁸

Interrater reliability in the pediatric population was found to have a kappa agreement of 1.0⁸. Interrater reliability was excellent in adults with peripheral nerve disorders (ICC = 0.91, 95%CI: 0.87, 0.94, $p < 0.05$).⁶ The intrarater reliability of six raters was good for four out of six raters (ICC 0.61 to 0.90); however, two raters had ICC values of 0.25 and 0.38.⁶

The TT was found to be better at detecting low levels of loss of sensation than the The Weinstein Enhanced Sensory Test (Bioinstruments Inc., Connecticut) (WEST), static, and moving two point discrimination tests in patients with carpal tunnel syndrome.⁹ Faught and McKee¹⁰ estimated that using a positivity criterion of sensation rated <10 on 3 of 4 of the first to fourth fingers was predictive of CTS (sensitivity = 80%, specificity = 48%, positive LR = 1.8). However, previous work identified that the TT was not sensitive in identifying asymptomatic sensation losses.¹¹

Several limitations of the TT should be noted. First, the Ten Test does not provide for the assessment of hyperesthesia, although some authors have suggested documenting hyperesthesia by scoring >10/10⁹ or altering the reference, using the 1-10 scale, where 1 = normal, and higher numbers represent increasing hyperesthesia,¹² although what 10/10 would represent is not clear.

Second, As a subjective test, rating the perception of sensibility may be challenging for some people.¹² A person may not be conscious of bilateral sensory loss, such as when sensory loss has been gradual and bilateral in an older population. Therefore, the comparison test site may not be normally innervated, as the test assumes, resulting in false negative outcomes.

Third, there has been limited research examining the clinical measurement properties of the Ten Test. Much of the data published has been reported by the developer. Therefore, further independent testing is needed.

Fourth, the clinical utility has been questioned. Clinical observation suggests many people have difficulty rating decreased sensation using a simple numeric rating scale of 1 to 10. The addition of verbal descriptors has been used in dyspnea and pain rating scales to provide clearer anchors and standardization of how patients calibrate their responses.^{2,3} The use of choices on a scale that are concrete versus abstract, with meaningful comparison between levels, and spontaneous responses may improve scale. Verbal scales may aid response consistency and be more intuitive than numerically ordered scales.²⁰ Analogue scales are more useful for statistical analysis.²⁰ Scales that hybridize visual analogue scales and verbal categories may combine the best of both features.²⁰ A combined visual, verbal and numeric version of the TT was created (see Appendix I). For the visual version, some modifications to the original version of the TT were made. First, the scale was changed to 0-10 as the authors felt that using 0 to represent no sensation was more intuitive than using a score of 1. Second, grey scale gradients were used to provide a visual effect of decreasing sensitivity. Third, descriptors were added to aid participants in describing and differentiating the different ratings.

The main purpose of this study was to examine the concurrent validity and patient preferences to this new, visual version of the TT. The visual version (VisTT) was expected to highly correlate with the original version as it is very similar, and both the TT and VisTT were expected to moderately correlate with vibration perception threshold

(VPT) and current perception threshold (CPT), as each tests sensation associated with A- β nerve fibres, albeit by different properties of the nervous system. The specific research questions being studied are:

1. What is the concurrent validity (Spearman's rank correlation coefficient r_s) for the visual version of the TT (VisTT), compared to the original TT, vibration perception threshold and current perception threshold tests in the population of people with decreased sensation in their fingers?

2. Do people with decreased hand sensation prefer the verbal or visual version the TT?

Methods

Ethics approval was received from the Hamilton Integrated Research Ethics Board. Ten subjects were recruited, each of whom reported decreased sensation in one or both hands. Participants were included if they were > 16 years of age, had decreased sensation in the hands/fingers, were capable of providing consent, and were able to participate in English. Participants were excluded if they had participated in a drug trial within the last 3 months. Recruitment was performed by public advertisement and from previous participation in research at the MacHand Lab, McMaster University.

Procedure

Participants attended a single one-hour session. The purpose and the potential risks were explained and demographic and consent forms were completed prior to completing the tests. A standardized examination of each of the tests was performed by a single rater (ND), who received training for the tests and performed practice applications to ensure competence. The testing order was: the TT, a randomized order of current perception threshold (CPT) and vibration perception threshold (VPT), followed by the visual version of the TT. Scores were recorded for each test, and the participants were then asked if they preferred the verbal or visual version. Testing was performed on the affected hand for unilateral cases, or the most affected hand for bilateral cases. CPT and VPT testing was performed on D1 (thumb), D2 (index finger) and D5 (little finger). All fingers were assessed with the TT, as per protocol, but values were recorded solely for D1, D2 and D5. Subjects were blinded to test results of all recorded scores; however, it was not possible to blind the rater.

The Ten Test.

The TT is an intensity rating scale of moving light touch threshold. The TT was performed according to the protocols proposed by the originator of the test.⁶

Participants were comfortably seated at a table and their forearms and hands were supported in a supinated position. The examiner applied moving light touch to an area perceived by the participant as normal. For participants with unilateral symptoms, the comparison site was the contralateral fingertip. In the case of bilateral symptoms, an alternate comparison site was used, often a site on the anterior, distal forearm. The chosen 'normal' area was defined as 10 on an analogue scale of 1 to 10 (1 = numbness, 10 = normal sensation). The affected fingertip was then simultaneously stroked and the participant was instructed to rate their sensibility compared to the control area.

Current perception threshold.

Current perception threshold testing (CPT) was performed with the Neurometer® (Neurotron, Inc, Baltimore, Maryland, http://www.neurotron.com/Neurometer_CPT-C.html). CPT uses electrical current to assess the function of sensory nerves.¹³ The Neurometer® provides a minimal electrical stimulus using an automated protocol.¹⁴ A sinusoidal constant alternating current is delivered at a choice of three different frequencies. These frequencies assess three types of afferent neurons: large myelinated A- β fibres (2000Hz), A- δ fibres (250Hz) and C fibres (5 Hz).¹⁵ Normal values are scores of 6-13, values < 6 are considered hypersensitive, and values > 13 demonstrate a loss of sensation.¹⁶

The participants were seated comfortably with their affected forearm and hand supported on a table. Gold-plated, 1cm diameter, gel-coated electrodes were applied to the medial and lateral distal fingertip and secured with adhesive tape. The procedure was explained, followed by a trial application, to allow participants to familiarize themselves with the test. Participants used a control button to initiate the onset of the delivery of stimulation and with which to respond at the moment that sensation was felt. The intensity of current applied during testing is not known until the test is completed. Only the 2000Hz frequency was applied, as this frequency most closely selects nerve fibres (large myelinated A- β fibres) that are associated with moving light touch threshold.¹⁷ The procedure was performed for the thumb, index and little finger to isolate median and ulnar nerve dermatomes. The CPT score for each finger was assessed with the range-CPT protocol, which reports the mean value after receiving three consistent responses.

Vibration perception threshold.

Vibration perception threshold testing (VPT) was assessed with the TSA-II NeuroSensory Analyzer (Medoc Ltd., Israel, <http://www.medoc-web.com/products/tsa-ii>). Participants were seated at a table with their affected forearm and hand supported on a pillow. The computer screen could not be seen by the participant. The assessment of VPT is standardized, although program parameters can be altered. We slowed the

ramp speed to reduce the error potential generated by response time delay. The method of limits was used, where an increasing intensity of vibration is conveyed through a hand-held applicator until the participant senses vibration, at which time they press a button on a computer mouse to indicate the VPT. A practice trial was performed, as there is decreased reliability in the first trial.¹⁸ Assessments were performed at the tip of the thumb, index and little finger. As per the equipment's protocol, eight trials were performed at each test site. The mean of the eight trials is reported as the VPT (microns).

The visual version of the TT.

The visual version of the TT was performed in the same method as the original TT, as described above, with the exception that a paper version was presented for the participant to use to report their rating. Visual TT scores were recorded. The participant was then asked, "Which version of the TT do you prefer, the verbal or visual format?", and the response was recorded.

Statistics

Concurrent Validity: All participants were evaluated with each of the four tests by one rater. Data was analyzed for each of the thumb (D1), index (D2) and little fingers (D5). Data was first plotted, and a Spearman's rank correlation coefficient (r_s) was calculated. Minimal significance level was set at $p < 0.05$. Strengths of Spearman's rank correlation coefficient (r_s) have been defined as ≤ 0.35 weak, $0.36 - 0.67$ moderate, and $0.68 - 1.0$ strong (Weber & Lamb, 1970; Mason, Lind & Marchal, 1983; as cited in Taylor, 1990).¹⁹ The percentage of preferences for each of the versions was calculated ($n = 10$). All analyses were performed using SPSS 20 for Windows (SPSS Inc., Chicago, IL).

Results

Participants. Fourteen (14) participants completed all testing. See Table 1 for a summary of participant statistics [mean, standard deviation (SD) and range (minimum, maximum)]. A wide age range was represented, and the diagnoses contributing to sensory loss were varied.

Sensory tests. Table 2 reports the median, mean and SD values for each test, as well as the coefficients of skewness and kurtosis. The mean and median scores of the Ten Test indicate that the participants sampled represented a group with perceived sensation loss on the mild end of the spectrum (median 8.5-9.5/10; mean = 7.64 - 7.79/10). The mean and median CPT values are within the normal score range.

The coefficients of skewness for the TT and VisTT scores indicate that the distributions range from symmetric to highly symmetric. TT and VisTT scores had a

negative skewness (range: -0.32 to - 1.17). TT and VisTT scores were not significantly different from that of a normal distribution. The coefficients of kurtosis for TT and VisTT scores were all negative, demonstrating that TT scores produced distribution curves that are lower and broader than that of a normal distribution. However, the peakedness of the distribution curves was not significantly different than that of a normal distribution.

Table 1. Participant Characteristics, n = 14			
Participant Characteristics			
Age (yrs)	<i>Mean</i>	<i>SD</i>	<i>Range</i>
	51.6	28.0	min 19, max 86
Sex (%)	<i>Male</i>	<i>Females</i>	
	n=2 (14%)	n=12 (86%)	
Symptoms (%)	<i>Unilateral</i>	<i>Bilateral</i>	
	6 (42.9%)	8 (57.1%)	
Most affected (%)	<i>Left</i>	<i>Right</i>	
	8 (57.1%)	6 (42.9%)	

SD- standard deviation, min - minimum, max - maximum

Table 2. Descriptive statistics n = 14					
Test	Median	Coefficient of skewness	Coefficient of kurtosis	Mean	SD
D1					
TT	9.50	-1.17	-0.10	7.64	3.25
CPT	13.00	-0.06	-1.19	13.07	7.64
VPT	0.77	1.87*	2.69*	0.96	0.85
Vis TT	8.50	-1.07	-1.07	6.79	3.38
D2					
TT	9.0	-0.70	-1.18	7.79	2.15
CPT	9.0	0.06	-1.14	9.07	6.06
VPT	0.52	1.91*	2.41*	1.08	1.30
Vis TT	7.50	-0.51	-0.51	7.00	2.45
D5					
TT	8.50	-0.82	-0.41	7.64	2.68
CPT	8.50	0.62	0.02	9.71	5.22
VPT	0.91	0.86	-0.11	1.07	0.68
Vis TT	7.50	-0.32	-0.32	7.57	2.24

*data is not normally distributed

TT = Ten Test

CPT = current perception threshold

VPT = vibration perception threshold

SD = standard deviation

D1 = thumb

D2 = index finger

D3 = little finger

Concurrent validity. The scores for the visual version of the TT correlated strongly with the original (verbal) version of the TT (D1 $r_s = 0.74$, $p = 0.000$; D2 $r_s = 0.90$,

$p = 0.000$; and D5 $r_s = 0.87, p = 0.001$). Correlation between the TT scores with CPT or VPT would be expected to produce negative values, as sensory loss is represented by lower TT scores and higher VPT and CPT scores. The relationship observed was negative, however, did not meet statistical significance with the exception of visual TT scores with VPT scores for D1 ($r = -0.54, p = 0.048$) (see Table 3). The Ten Test does not appear to correlate with CPT or VPT scores. Due to the small sample size in this study and the lack of variation between subjects, the results for the Visual Ten Test are equivocal.

Table 3. Correlations between the TT, current perception threshold and vibration perception threshold.

Test	D1			D2			D5		
	TT	CPT	VPT	TT	CPT	VPT	TT	CPT	VPT
TT	-	-0.30 (0.30)	-0.40 (0.16)	-	-0.18 (0.53)	-0.18 (0.55)	-	-0.03 (0.92)	-0.02 (0.99)
Visual TT	0.74* (0.002)	-0.38 (0.18)	-0.54* (0.048)	0.90* (0.000)	-0.16 (0.58)	-0.32 (0.26)	0.87* (0.000)	-0.01 (0.98)	-0.02 (0.96)

Note: Correlation values represent Spearman’s rank correlation coefficient r_s values with p values represented in brackets

* significant p value (<0.05)

TT = Ten Test

CPT = current perception threshold

VPT = vibration perception threshold

D1 = thumb

D2 = index finger

D3 = little finger

Participant preference: Twelve out of fourteen (85.7%) participants preferred the visual version of the TT (See table 4).

Table 4. Participant preferences of TT format (n=14)

Preference	Frequency	Percent %
Verbal	2	14.3
Visual	12	85.7
Total	14	100

Discussion

The visual version of the Ten Test correlates highly with the verbal version and is preferred by most people in an adult population with decreased sensation in the hand. Therefore, the VisTT version is an acceptable alternate format to provide to patients.

Overall, correlation was not observed between the scores of either version of the TT and CPT or VPT scores, with the exception of the VisTT and VPT for D1. However, the fact that the correlations to CPT and VPT were similar between the verbal and visual versions of the TT provides support that the alternate forms of the TT are measuring in the same way. Overall, TT test scores do not appear to be able to predict CPT or VPT scores. The Visual Ten Test may relate better to other sensory tests. Both versions of the TT correlated similarly to CPT and VPT scores.

A possible explanation for the lack of association between the VisTT and CPT and VPT scores may be related to the differences between the structures stimulated by each test. CPT directly stimulates nerve fibres, since the electrical current intensities provided are below that which would stimulate sensory receptors.⁴ VPT activates sensory receptors directly. The TT depends on a person's perception of sensation; this is reliant upon what a person perceives as "normal" and their ability to quantify that perception. Thus, the lack of correlation between the TTs with CPT and VPT may be due, in part, to the tests assessing different components of the sensory pathways.

We made a number of changes to the TT: adding verbal descriptors, creating a graphic image to describe the scale and the descriptors, and adding a zero to the scale. There were potential implications to all of these changes. Typically there are reasons not to change a scale as originally described by its developer since the data published on the scale would be no longer directly comparable. However, as the TT is not yet widely used, any improvements to its validity and acceptability for clinical practice should be made before widespread usage has occurred and substantial pools of comparative data are available. Thus, this is an ideal time to make these changes if they improve the measure. Further, the changes that we are recommending have already been informally reported by others as a means of improving the validity of the TT (<http://nervesurgery.wustl.edu/ev/evaluation/sensory-specificexam/Pages/sensorytests.aspx>). Adding a zero to the scale was thought to enhance clarity since it represents a complete absence of sensation. This could potentially lower scores in patients with no sensation, since that number is now anchored lower. However, absent sensation is a rare finding unless a nerve has been transected and in populations with mild sensory disturbances such as in this study the estimates are unlikely to be changed.

Participants strongly favoured the visual version of the TT (85.7%). Although qualitative questions were not formally asked to delineate why they preferred the visual version, several participants voluntarily described that it was useful to have descriptors to assist in rating sensibility and that they preferred to be able to "look at" the test. Providing information in verbal and visual formats can enhance clarity by tapping into different cognitive pathways. Further, the grayscale gradient may enhance the concept

that sensation exists on a continuum. While this approach may not address situations where patients have a hard time understanding the concept of what sensation is, greater focus on describing the concept being measured with clarity should improve measurement validity.

The primary limitation of this study is the small sample size, resulting in decreased precision in the estimates of correlation coefficients and greater potential for lack of representative subjects. The assessor was not blinded to the results of the TT and was responsible for recording test results which may have contributed observer bias. As participants operated the VPT and CPT controls, scores were not influenced by the examiner.

The results from this study indicate that the visual version of the Ten test may be an improvement of the original version. The original version has demonstrated strong reliability and strong correlation to the SWMT. As the visual version correlates highly with the original version good clinical measurement properties for the visual version are predicted. The visual version demonstrated better correlation to VPT and this may be due to the use of descriptive anchors better standardizing responses.² The visual version was preferred and may be less challenging to implement in a clinical setting. The visual version of the Ten test should be further explored as it may exhibit better reliability, validity and clinical utility. However, the results of the TT are based on a preliminary investigation with a small sample and are insufficient to be confident that the visual version is conceptually and statistically superior to the verbal version. Cognitive interviewing, reliability and validity testing in larger samples and examination of responsiveness are needed to provide sufficient evidence to indicate superiority, if present.

Conclusion:

The visual version of the TT correlates highly with the original version, which has been found to be reliable and valid. Neither TT version correlated with VPT or CPT scores; but demonstrated similar correlations regardless of whether the verbal or graphic approach was used. Participants strongly preferred the visual version over the verbal version. Being a simple and reliable tool, the visual version of the TT would be beneficial for use in clinical practice and is an acceptable alternate format to provide to patients.

Appendix I. The Visual Ten Test.

Visual Ten Test	
10	Normal
9	Almost the same
8	
7	Fairly good
6	
5	Half of my normal
4	A little bit
3	
2	Almost none
1	
0	Can't feel anything

Instructions:

Please rate how much sensation you feel compared to a normal part of your body.

References

1. Atroshi I, Gummesson C, Johnsson R, Ornstein E, Ranstam J, Rosen I. Prevalence of carpal tunnel syndrome in a general population. *The Journal of the American Medical Association*. 1999;282(2):153-158.
2. Warwick D, Dunn R, Melikyan E, Vadher J. Assessment. *Hand Surgery: Oxford Specialist Handbooks in Surgery*; 2009:1-37.
3. Jerosch-Herold C. A study of the relative responsiveness of five sensibility tests for assessment of recovery after median nerve injury and repair. *The Journal of Hand Surgery: British & European Volume*. 6 2003;28(3):255-260.
4. Bear-Lehman J, Abreau B. Evaluating the hand: issues in reliability and validity. *Physical Therapy*. 1989;69(12):11.
5. Jerosch-Herold C. Assessment of sensibility after nerve injury and repair: a systematic review of evidence for validity, reliability and responsiveness of tests. *The Journal of Hand Surgery: British & European Volume*. 6 2005;30(3):252-264.
6. Strauch B, Lang A, Ferder M, Keyes-Ford M, Freeman K, Newstein D. The Ten Test. *Plastic and reconstructive surgery*. 1997;99(4):1074-1078.
7. Strauch B, Lang A. The Ten Test revisited. *Plastic & Reconstructive Surgery*. 2003;112(2):593-594.
8. Sun HH, Oswald TM, Sachanandani NS, Borschel GH. The Ten Test: Application and limitations in assessing sensory function in the paediatric hand. *Journal of Plastic, Reconstructive & Aesthetic Surgery*. 2010;63(11):1849-1852.
9. Patel MR, Bassini L. A comparison of five tests for determining hand sensibility. *Journal of reconstructive microsurgery*. 1999;15(7):4.
10. Faught BE, McKee NH. Establishing a positivity criterion in determining the utility of the Ten Test in diagnosing carpal tunnel syndrome. *Journal of reconstructive microsurgery*. 2002;18(7):634.
11. Durrant N. *The reliability and validity of the Ten Test and the validity and patient preference of a visual version of the Ten Test*. Hamilton: Rehabilitation Science, McMaster University; 2014.
12. Uddin Z, MacDermid J, Packham T. The Ten Test for sensation. *Journal of Physiotherapy*. 2013;59(2):132-132.
13. Bleecker ML, Katims JJ, Patil AS, et al. Current perception threshold screening for carpal tunnel syndrome. *Archives of Environmental Health*. Vol 46; 1991:207+.
14. MacDermid JC, Gross A, Galea V, et al. Developing biologically-based assessment tools for physical therapy management of neck pain. *Journal of Orthopaedic & Sports Physical Therapy*. 2009;39(5):14.
15. Lowenstein L, Jesse K, Kenton K. Comparison of perception threshold testing and thermal-vibratory testing. *Muscle & nerve*. April 2008 2008;37(4):514-517.
16. Neurometer CPT/C Manual. *Operating Manual* Denver, Colorado: Neurotron, Incorporated; 2010:93.

17. Dotson RM. Clinical neurophysiology laboratory tests to assess the nociceptive system in humans. *Journal of clinical neurophysiology*. 1997;14(1):32-45.
18. Grunert BK, Wertsch JJ, Matloub HS, McCallum-Burke S. Reliability of sensory threshold measurement using a digital vibrogram. *Journal of Occupational Medicine*. 1990;32(2):100-102.
19. Taylor R. Interpretation of the correlation coefficient: A basic review. *Journal of diagnostic medical sonography*. 1990;6(1):5.
20. Lansing RW, Moosavi SH, Banzett RB. Measurement of dyspnea: word labeled visual analog scale vs. verbal ordinal scale. *Respiratory Physiology & Neurobiology*. 2003;134(2):7.
21. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain*. 1986;27(1):10.

Chapter Three: Future Directions

Summary.

The use of sensory evaluation of the hand in clinical practice is limited. Two contributing factors are: 1) the time and economic burden of sensory testing and, 2) a deficit of information about the clinical measurement properties of sensibility measures. The two studies contained in this thesis attempted to advance the knowledge about the clinical measurement properties of the Ten Test, current perception threshold, and vibration perception threshold tests in the population of people with hand sensory deficits. In particular, the Ten Test has low burden of time and no financial cost. A visual version of the Ten Test is an alternate version of the Ten Test that shows early promise of validity and patient preferences and it worthy of further validation.

Limitations.

Several limitations of the research conducted have already been described in the two studies contained within this thesis, the most significant being small sample sizes.

Key Messages.

There is a lack of information available to clinicians to allow the interpretation of sensation test scores on individual patients. The first study has contributed minimal detectable change scores (MDC_{90}) for the Ten Test, VPT and CPT with which a clinician can judge whether a true change in health status has occurred in a person.

The first study has demonstrated that the Ten Test is a reliable tool for people with decreased hand sensation. However, as it is based on perception, the Ten Test may fail to identify asymptomatic losses of sensitivity. Nonetheless, clinicians can now be more confident in use of the Ten Test as an outcome measure for clients with hypoesthesia, while being aware of its limitations.

VPT is more likely to be used for research purposes due to the burden of time to implement testing and the high cost of equipment. VPT scores have been shown to be reliable; however, CPT scores have less reliability and do not appear to add significantly useful information beyond VPT scores.

The second study examined a newly created visual version of the Ten Test. It was found to highly correlate with the original version of the Ten Test, and was preferred by participants in a small sample.

Future Recommendations.

Further evaluation of the clinical measurement properties of the Ten Test with larger sample sizes, a wider spectrum of sensory problems and over time is needed. Perhaps most critically lacking are the examination of test-retest reliability assessed over several days, independent testing for inter-rater and intra-rater reliability in the adult population, and responsiveness. Also, the Ten Test needs to be evaluated in its ability to detect asymptomatic sensation deficits.

The visual version of the Ten Test is a promising tool and warrants further clinical measurement testing especially in light of the strong patient preference for the visual over

the verbal version. It may enhance validity, but whether this is through increased comprehensive would require cognitive interviewing tests. Its performance in relation to the standard TT and other sensory measures in terms of detecting clinical change is an important next step.

Lastly, research is not useful if not known. The Ten Test appears to be a useful clinical tool that would likely be readily accepted by clinicians. Yet, it is not widely used. Therefore, knowledge translation strategies are required to further the awareness of the Ten Test and its clinical measurement properties to improve the ability of clinicians to help clients to the best of our abilities.