ASSESSMENT OF SENSIBILITY AFTER NERVE INJURY AND REPAIR: A SYSTEMATIC REVIEW OF EVIDENCE FOR VALIDITY, RELIABILITY AND RESPONSIVENESS OF TESTS

C. JEROSCH-HEROLD

From the School of Allied Health Professions, University of East Anglia, Norwich NR4 7TJ, UK

Recovery of sensibility after peripheral nerve injury and repair needs to be assessed using psychometrically robust measures. In this study the literature was reviewed to identify what tests are available to quantify sensibility and to assess their validity, reliability and responsiveness. The databases Medline, CINAHL, Embase and AMED were searched for studies reporting the psychometric properties of sensibility tests. While there is a plethora of tests and studies reporting the outcomes after peripheral nerve suture only a few of the tests have evidence of validity, reliability and responsiveness. Currently the touch threshold test using monofilaments such as the Weinstein Enhanced Sensory Test (WEST) or Semmes–Weinstein Monofilament Test (SWMT) and the shape–texture identification (STI™) test for tactile gnosis are the only tests which meet criteria for a standardized test and have had their psychometric properties evaluated and quantified. Journal of Hand Surgery (British and European Volume, 2004)

Journal of Hand Surgery (British and European Volume, 2005) 30B: 3: 252–264

Keywords: sensibility, nerve injury, outcome measures, systematic review

INTRODUCTION

Injuries to peripheral nerves in the upper extremity are a common cause of morbidity and disability (Jacquet et al., 2001). Despite surgical advances, the return of sensory and motor functions is poor in adults and can take up to 7 years (Rosén et al., 2000a). Patients also suffer cold intolerance, pain and paraesthesiae (Rosén and Lundborg, 2000b; Sunderland, 1978).

Proponents of new biological (McKay Hart et al., 2003) and surgical techniques (Lundborg et al., 2004) for the acute management of injured nerves, and those advocating novel ways of re-educating sensibility (Rosén et al., 2003) have also identified the need to test these interventions in well-designed prospective randomized controlled trials. The use of outcome measures which are robust in capturing meaningful and important changes over time and between groups is paramount. Furthermore it demands the use of instruments and protocols for administration and scoring which are standardized and therefore can be used and compared between different centres.

Instruments that are used for research and the documentation of clinical outcomes need to be precise in the neurophysiological and functional parameters which they purport to test, yet must also be able to detect a wide range of deficit from complete anaesthesia to near normal sensibility.

Given the long-term follow-up of patients with peripheral nerve injury, instruments and methods for testing hand sensibility are required which capture meaningful information, are reproducible over time and responsive to small but clinically important changes. These psychometric properties are defined in measurement theory as test validity, test reliability and test responsiveness, respectively, and determine the degree of confidence that the clinician, patient or researcher can have in the conclusions drawn from the data. The clinical utility, that is the acceptability, practicality and cost of the measures also needs to be borne in mind. This review aims to critically evaluate the evidence on the validity, reliability and responsiveness of tests for hand sensibility in peripheral nerve injury and repair. The terms validity, reliability and responsiveness are well defined in the methodological literature (Streiner and Norman, 2003) and a full discussion of these concepts is beyond the scope of this review. However, it must be emphasized that these attributes are not fixed but have to be assessed within the context of a specific population, settings and instruments used.

CLASSIFICATION OF SENSIBILITY TESTS – AN UNRESOLVED PROBLEM

The touch function of the hand defies simple description, as is evident from the many different instruments, developed to quantify sensibility. At the simplest level, sensibility can be divided into protective and discriminative sensibility, the latter being the functional sensibility which enables humans to resolve fine spatial detail such as texture or shape. It is this function which often returns incompletely and which is the focus of this review. Despite the plethora of tests that exists, there is little consensus on what tests should be used in outcome studies.

One approach is to group the different tests according to the specific receptor population which they purport to

test. Such a classification system was first proposed by Dellon (1981) in which sensory tests were correlated with four types of mechanoreceptive units found in the glabrous skin of the hand. Initially, this may appear a well-reasoned approach. However, a probe such as a monofilament is likely to engage fast- and slow-adapting receptors due to the vibration generated by the examiner when the filament is held in place and the dynamic deformation of the skin during placement and withdrawal of the filament (Bell-Krotoski, 1989). It, thus, cannot be classed as a "static" test or one which only measures the slowly adapting receptor population. Thus, the classification of tests according to the specific receptor populations which they are purported to assess is misleading.

Perhaps the most useful classification available to date is that proposed by Fess (1995), who classified tests according to a hierarchy: the first group is composed of tests which assess detection thresholds for light touch, deep pressure and dynamic stimuli such as vibration. These tests address the question of "can you feel the stimulus or not?", and allow quantification of the lowest threshold at which detection occurs. At the next level are tests of spatial discrimination. These include tests which assess the smallest spatial threshold at which localization or discrimination between different stimuli occurs, for example distinguishing between 1 or 2 calliper points, the orientation of gratings or the distance between the actual and the perceived stimulus. The third category comprises tests which require identification of objects, shapes or textures, that is the micro- or macrogeometry of objects. These often require active movement to resolve the spatial aspects of the items.

REVIEW OBJECTIVES

The purpose of this paper is to present a systematic review of studies evaluating the psychometric properties of sensibility tests and to provide a rationale for the selection of tests when assessing outcome in peripheral nerve injuries. The degree of deficit as well as the processes involved in repair and recovery in these patients are different from those with nerve compression or neuropathies and therefore only studies including subjects with peripheral nerve injuries are included.

METHODS

A database search was undertaken on *Medline* (1966–2004); *CINAHL* (1982–2004); *Embase* (1980–2004) and *AMED* (1985–2004). The following keywords were used alone or in combination and mapped to MeSH headings: sensation, touch, instrumentation, peripheral nerves, nerve injuries, outcome assessment, upper extremity, adults, humans. In order to maximize retrieval of all relevant studies, the Journal of Hand Therapy (from 1989), British Journal of Hand Therapy (from 1988) and Journal of Hand Surgery were hand searched for relevant articles, as were the reference lists of those articles already obtained. Only English language articles were included. The criteria for inclusion of articles in this review were: (i) study designed with the primary purpose of evaluating aspects of validity, reliability or responsiveness of sensibility instrument(s); (ii) study sample included subjects who had undergone peripheral nerve injury and repair in the hand or upper limb. Studies based solely on healthy controls or patients with nerve compression were excluded, as were purely descriptive studies reporting outcomes.

Data on the study design, subjects, instruments and results were extracted and tabulated for each study by the author including the statistical measures of validity, reliability and responsiveness where these were available. Meta-analysis of the results was not possible. These are presented as a summary in Table 1. Using the classification of tests presented earlier, the evidence on validity, reliability and responsiveness for each of the tests identified were tabulated to offer a synthesis of evidence on the psychometric properties for each sensibility test. These are presented in Tables 2–4.

RESULTS

A total of 15 studies were included, several of these were written by the same authors. Several of the articles evaluated more than one test and examined several psychometric properties of the tests. The sensibility tests identified and studied were: touch threshold tests using monofilaments (SWMT and WEST), vibration thresholds with tuning forks and vibrometers, static and moving two-point discrimination, localization tests and tests of object, shape and texture identification. Each study was reviewed by the author with regards to the validity of the methods used and the findings pooled to provide a comprehensive summary of the evidence regarding reliability and validity, and responsiveness for each group of tests. These are summarized in Tables 2–4 under (i) measures of touch detection, (ii) measures of spatial discrimination and (iii) measures of tactile gnosis. A critical evaluation of the quality and quantity of this evidence base is presented in the discussion and forms the basis for recommendations for practice and research.

DISCUSSION

Methodological considerations of studies

A large number of studies were excluded because the reporting of an aspect of validity or reliability was secondary to the presentation of outcomes. In other words these studies were not designed with the primary