Prospective Analysis of Relationships of Outcome Measures for Ulnar Neuropathy at the Elbow

Rajiv Midha, Jason Noble, Vivek Patel, Peter H. Ho, Catherine A. Munro, John Paul Szalai

ABSTRACT: Background: We undertook a prospective study to investigate relationships between outcome measures of ulnar neuropathy at the elbow. Methods: Thirty-one patients (mean age 52.6, range 20-80), with clinically and electrically verified ulnar neuropathy at the elbow, were seen independently by a neurosurgeon and a physiotherapist. All tests were administered to all patients on each visit. Data collected included measures of sensory (monofilament, two-point discrimination, vibration) and motor function (grip, key-pincho, muscle atrophy), pain (visual analogue scale (VAS)) and impact on lifestyle (Levine’s questionnaires (function status score – FSS, symptom severity score – SSS)), disability of the arm, shoulder and hand module (DASH) and patient-specific measures (PSM). Parametric and non-parametric correlation and factor analysis were done. Results: Outcome analysis was available for 63 patient visits, with follow-up obtained for 20 patients (mean 8.5 months). Lifestyle and pain instruments (FSS, SSS, DASH, PSM and VAS) all correlated well with each other (r > 0.6, p < .01). DASH was moderately to highly correlated to nine of the 11 measures. Some tests correlated poorly, for example, Semmes-Weinstein monofilament with other sensory measures and muscle atrophy with almost all measures. Factor analysis revealed that there are two principal factors, accounting for 77% of the variance. Factor 1 relates to impact on lifestyle and pain while Factor 2 relates to strength and function. Discussion/Conclusions: Intraclass measures, particularly ones assessing lifestyle and pain instruments are strongly correlated. Factor analysis revealed two principal factors that account for the majority of the variance; future studies with a larger sample size are needed to validate this analysis.

RÉSUMÉ: Analyse prospective des relations entre les mesures de résultats dans la neuropathie cubitale au niveau du coude. Introduction: Nous avons effectué une étude prospective afin d’étudier les relations entre les mesures des résultats dans la neuropathie cubitale au niveau du coude. Méthodes: Trente et un patients (âge moyen 52.6 ans, âge de 20 à 80 ans), atteints de neuropathie cubitale au niveau du coude, vérifiée cliniquement et électriquement, ont été examinés indépendamment par un neurochirurgien et un physiothérapeute. Tous les tests ont été administrés à tous les patients à chaque visite. Les données incluent des mesures de la fonction sensitive (monofilament, discrimination spatiale, vibrations) et motrice (préhension, atrophie musculaire), de la douleur (échelle analogique visuelle (VAS)) et impact sur le style de vie (questionnaires de Levine (score de statut fonctionnel –FSS, score de sévérité des symptômes – SSS)), invalidité du bras, du module épaule et main (DASH) et mesures spécifiques du patient (PSM). On a procédé à une analyse de corrélation et à des analyses factorielles paramétriques et non paramétriques. Résultats: L’analyse des résultats était disponible concernant 63 visites de patients, dont des visites de suivi chez 20 patients (moyenne de 8.5 mois). Les mesures se rapportant au style de vie et à la douleur (FSS, SSS, DASH, PSM et VAS) étaient bien corrélées entre elles (r > 0.6, p < 0.01). Le DASH était de modérément à fortement corrélé à neuf des 11 mesures. Certains tests étaient peu corrélés, tels le test du monofilament de Semmes-Weinstein aux autres mesures sensitives et l’atrophie musculaire à presque toutes les mesures. L’analyse factorielle a montré qu’il y a deux facteurs principaux qui expliquent 77% de la variance. Le premier facteur est associé à l’impact sur le style de vie et la douleur, alors que le deuxième facteur est associé à la force et à la fonction. Discussion/Conclusions: les mesures intraclasse, particulièrement celles qui évaluent le style de vie et les mesures de la douleur sont fortement corrélées. L’analyse factorielle a montré deux facteurs principaux qui expliquent la plus grande partie de la variance; d’autres études portant sur un échantillon plus considérable seront nécessaires pour valider les résultats de cette analyse.

neuropathy at the elbow. These tests assess sensory and motor function, as well as pain and impact on lifestyle. No single test can be relied upon to effectively demonstrate a patient’s actual clinical condition. In the absence of a gold standard clinical test, therapists and physicians must administer a number of the available measurement tools to correctly reflect and track a patient’s progress. Classification systems to date use motor and sensory criteria to categorise patients into mild, moderate and severe cases, but have not been validated. Instruments that measure patient satisfaction and functional assessments have not been taken into consideration.

The current literature regarding the clinical outcome of ulnar neuropathy at the elbow is deficient in that studies have examined relations among only a limited number of the available assessment tools. For instance, attempts have been made to discover correlation between specific outcome measures, but no prospective study has yet examined a substantial number of assessment tools with the objective of exploring relationships between the tests and determining which tests are sensitive to clinical change and predictive of clinical outcome.

In light of the overabundance of outcome measures, we undertook to examine relationships among the various outcome measures, using correlation statistics. We also sought to determine which instruments predict clinical status in subsequent visits and to ascertain how well the diagnostic tools capture the trend in a patient’s clinical status. The long-term goal is to identify a unique combination of outcome measures which accurately depicts clinical status, in order to develop a concise ulnar neuropathy battery. We report the first step in developing a battery, using an exploratory factor analysis to extract related and covarying factors from among the outcome variables. The present study is unique in that it is prospective in nature, a substantial number of the available outcome measures were administered to each patient at each visit, and relationships both within and between patients have been investigated using correlation and factor analysis.

**Patients and methods**

Thirty-one patients with a mean age of 52.6 (range 20-80; 25 males, 6 females) were seen independently by a neurosurgeon (R.M.) and a physiotherapist (P.H.) at Sunnybrook & Women’s College Health Sciences Centre between 1996 and 1999. All patients were known to have some degree of ulnar neuropathy based on a combination of clinical and confirmatory electrodiagnostic findings, consisting of a localized conduction block or at least a 10 m/s slowing of ulnar nerve conduction affecting the elbow segment. Nineteen patients had surgical intervention. Patients presenting with symptoms of bilateral ulnar neuropathies and of carpal tunnel syndrome were excluded from the study.

During each visit, a standard ulnar nerve assessment was administered (Table 1). This assessment included sensory, motor and functional tests. Sensory parameters assessed were static and moving two point discrimination (s2-PD, m2-PD), tuning fork vibration tests at 265 cps (VIB), and Semmes-Weinstein monofilament testing. All sensory tests were applied to the ulnar innervated fifth digit to assess dysfunction, as well as patients’ second digit (median nerve innervated) as a control. Motor examination consisted of a standard Medical Research Council assessment of ulnar innervated muscles, grip strength, key pinch strength, and noting the presence of any muscle atrophy in ulnar innervated hand intrinsics, namely the first dorsal interosseous and hypothenar muscles. Muscle strength measures were recorded using a Jamar dynamometer.

The functional and pain tests administered included visual analogue scale (VAS), Levine’s functional status (FSS) and symptom severity scales (SSS),7 the disability of arm, shoulder and hand core module (DASH),10 as well as patient specific measures (PSM).11 Patient specific measures is a subjective measure where patients identify up to five activities they have difficulty with as a result of their neuropathy, and then grade their ability to perform the activities on an 11 point scale. As a substantial number of patients were unable to report five PSM, the mean of only the first three reported scores was considered in the statistical analysis.

Since it is known that comparing muscle strength of uninjured areas can be sensitive in evaluating nerve regeneration,8 patients’ contralateral nonsymptomatic arms were assessed, and, accordingly, all muscular measures were reported relative to the uninjured arm. Key pinch and grip strength measures were analysed as a ratio of the force recorded for the symptomatic arm over the force measured by the nonsymptomatic arm. In order to account for the inherently greater strength of the dominant arm, a correction factor of 10% was employed.

Data were collected prospectively and entered into a common spreadsheet. Dichotomous coding was used to report certain outcome measures, in keeping with accepted practice.9,12,13 Any presence of muscle atrophy was considered abnormal (score = 1). Monofilament scores greater than 2.83 were considered abnormal (score = 1), and values below this conventional threshold were considered normal (score = 0).9 For both s2-PD

<table>
<thead>
<tr>
<th>Test</th>
<th>Abbreviation</th>
<th>Nature</th>
<th>Normal range/ Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monofilament</td>
<td>Mono</td>
<td>Sensory</td>
<td>≤ 2.83</td>
</tr>
<tr>
<td>Static two-point</td>
<td>s2-PD</td>
<td>Sensory</td>
<td>&lt; 5 mm</td>
</tr>
<tr>
<td>discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moving two-point</td>
<td>m2-PD</td>
<td>Sensory</td>
<td>&lt; 5 mm</td>
</tr>
<tr>
<td>discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vibration test at 265 cps</td>
<td>VIB</td>
<td>Sensory</td>
<td>0 (no paraesthesia)</td>
</tr>
<tr>
<td>Grip ratio</td>
<td>Grip</td>
<td>Motor</td>
<td>1.0</td>
</tr>
<tr>
<td>Key pinch ratio</td>
<td>Key</td>
<td>Motor</td>
<td>1.0</td>
</tr>
<tr>
<td>Muscle atrophy</td>
<td>MA</td>
<td>Motor</td>
<td>0</td>
</tr>
<tr>
<td>Patient specific measures</td>
<td>PSM</td>
<td>Lifestyle</td>
<td>10.0</td>
</tr>
<tr>
<td>Functional status scale</td>
<td>FSS</td>
<td>Lifestyle</td>
<td>1.0</td>
</tr>
<tr>
<td>Symptom severity scale</td>
<td>SSS</td>
<td>Lifestyle/Pain</td>
<td>1.0</td>
</tr>
<tr>
<td>Disability of the arm, shoulder and hand core module</td>
<td>DASH</td>
<td>Lifestyle</td>
<td>0</td>
</tr>
<tr>
<td>Visual analogue scale</td>
<td>VAS</td>
<td>Pain</td>
<td>0</td>
</tr>
</tbody>
</table>
and m2-PD, scores greater than or equal to 5 mm were considered abnormal (score = 1), while values below 5 mm were deemed normal. Both hypersensitivity and diminished perception in vibratory sensation were considered abnormal (score = 1).

Data from the first visit provided the primary focus of statistical analyses, although parallel analyses were executed on the data from the second visit. Relationships between the scores on individual tests or responses were analysed through the appropriate correlational statistics: Phi, point biserial or Pearson product moment correlations. For the first visit dichotomous data, Cronbach’s α internal consistency analysis was applied to the summary scores. For the continuous measures and the summarised dichotomous scores, an exploratory factor analysis was performed to extract related and covarying factors. Specifically, we used a principal component analysis to extract factors and the varimax (oblique) method for rotation. To test the expected improvement in the measures from visit 1 to visit 2, paired t-test on the factor scores was applied. Note that factor scores for the second visit were generated from factor coefficients of the first visit. Statistical analysis was done using SPSS 10.0 software for PC.

Results

Sixty-three individual assessments were made. Follow-up assessments were obtained for 20 patients (mean follow-up 8.5 months, range 3-24 months), and seven patients had a third follow-up visit. Follow-up could not be obtained for eleven patients, however three of these patients responded to mail-in questionnaires. While no single reason for patients failure to return for follow-up could be identified, a number of these patients had moved with no forwarding address given.

Correlation analysis

Correlation analysis revealed several significant correlations. Within the first visit, the DASH score was moderately to highly correlated with nine of the 11 analysed tests (r > 0.440, p ≤ 0.030). This excluded only muscle atrophy and monofilament sensory testing. DASH did not correlate as significantly in the second visit (only with five of the 11 measures) as it did in the first visit. Similarly, s2-PD correlated significantly with many measures in the first visit (eight of 11 measures), but not in the second visit. SSS was unique in that it correlated well with many of the measures in both the first (with eight of 11) and second (seven of 11) visits (r > 0.300, p ≤ 0.050).

Measures which assess pain and impact on lifestyle (FSS, SSS, DASH, PSM, VAS) all correlated well with one another (r > 0.60, p ≤ 0.010) within both visits. These measures also correlated well with many of the motor and sensory tests.

Monofilament sensory testing correlated most poorly and the least significantly of all the analysed outcome measures, only correlating significantly with SSS in the first visit and VIB in the second visit. The other sensory tests (VIB, s2-PD, m2-PD) correlated significantly with one another (r > 0.370, p ≤ 0.035) as well as with many of the other outcome measures in the first visit.

The three motor/muscle tests (muscle atrophy, grip strength, key pinch) correlated well with one another (r > 0.40, p ≤ 0.020) within the first visit. Although grip strength and key pinch did have some significant correlations with some of the other measures, muscle atrophy did not correlate significantly with any of the non-motor tests in any of the analyses.

Factor analysis using principal component method

Cronbach’s alpha analysis of the five dichotomous variables revealed an alpha value of 0.667. Coefficient alpha increased to a value of 0.737 when muscle atrophy was deleted from the scale, indicating that muscle atrophy is inconsistent with the other (all sensory) dichotomous measures. The significant correlations between the dichotomous outcome measures and the internal consistency (excluding muscle atrophy) indicated by Cronbach’s alpha suggests that there is one common factor which underlies these outcome measures. Thus, the dichotomous variables were grouped together into one variable, which we termed the combined sensory score, and used in the subsequent factor analysis.

Factor analysis revealed that there is one principal factor on which all eight variables (DASH, FSS, SSS, VAS, key pinch, grip strength, PSM and the combined sensory score) are loaded (Table 2). With an eigenvalue of 5.388, this one factor accounts for 59.86% of the variance. A second principal factor accounts for 16.85% of the variance, with an eigenvalue of 1.516. Together these two factors account for 76.71% of the variance. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.755, suggesting that factor analysis was appropriate. Component score coefficients are given in Table 2. As the variables loading most strongly on Factor 1 related to impact on lifestyle and pain, we have labelled this factor ‘Lifestyle/Pain’. Variables loading most strongly on Factor 2 are objective measures of strength and function, and we have labelled this factor ‘Function’.

Table 2: Factor score coefficient matrix revealed by factor analysis

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Component Coefficient</th>
<th>Descriptive Statistics Visit 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lifestyle/ Pain Factor</td>
<td>Function Factor</td>
</tr>
<tr>
<td>Grip</td>
<td>0.011</td>
<td>0.290</td>
</tr>
<tr>
<td>Key</td>
<td>0.028</td>
<td>0.288</td>
</tr>
<tr>
<td>VAS</td>
<td>0.253</td>
<td>0.018</td>
</tr>
<tr>
<td>FSS</td>
<td>0.194</td>
<td>-0.078</td>
</tr>
<tr>
<td>SSS</td>
<td>0.284</td>
<td>0.058</td>
</tr>
<tr>
<td>DASH</td>
<td>0.187</td>
<td>-0.117</td>
</tr>
<tr>
<td>PSM</td>
<td>-0.069</td>
<td>0.193</td>
</tr>
<tr>
<td>Combined Sensory Score</td>
<td>0.207</td>
<td>0.042</td>
</tr>
</tbody>
</table>

Factor scores for new patients can be calculated using the following equation: \( FS = \sum (component\ coefficient \times Z\-score) \), where: \( z\-score = (individual's\ score - mean\ value) / standard\ deviation \).
Outcome measures over time

The majority of the patients (whether operated on or not) improved clinically on subsequent visits (Figure). The various outcome measures, particularly those related to lifestyle and pain, demonstrated striking concordance to each other (Figure). Indeed, paired t-test on the factor scores suggested a significant improvement in factor scores from visit 1 to 2 (p = 0.033), supporting the construct validity of the proposed model. However, the true construct validity can only be determined by analysing a second sample of patients, with factor scores for new patients calculated using the equation presented in Table 2. Such analysis might include assessment of surgical success.

DISCUSSION

There remains no consensus as to which surgical treatment for ulnar neuropathy at the elbow is superior. A considerable problem is the numerous outcome measures used by different authors in deeming the failure or success of their intervention. Indeed, even retrospective comparison of different studies is hampered by the variability of outcome reporting. Treatment trials should rely on an accurate, reproducible and valid outcome system or scale to determine if one treatment is superior over others. As there have been calls for a standard classification system, establishing an accurate and standardised assessment protocol for ulnar neuropathy can allow this to be examined.

Factor analysis is a statistical technique that is used to identify a relatively small number of factors that can be used to represent relationships among sets of many interrelated variables. The basic assumption of factor analysis is that the factors underlying complex phenomena can be used to explain the phenomena itself. This is done by identifying indirectly observable factors based on a set of observable variables. Ideally, factor analysis should be able to explain observed correlations using a small number of factors. In the present report, we have detected two factors, which underlie the various instruments used in the assessment of ulnar nerve entrapment neuropathy. We have termed these factors Lifestyle/Pain and Function. These factors should now be explored using a larger set of data to confirm that this analysis is correct and to further define the factor. This should allow for the development of an assessment battery, using fewer instruments.

Measures within the same diagnostic realm (i.e. sensory, motor, pain or lifestyle) correlated well with one another, indicating that measures are internally consistent. Lifestyle/Pain instruments were also found to be strongly correlated to each other in the first and in subsequent visits. The correlation data also suggest that there is considerable redundancy in the outcome instruments. On the other hand, measures that assess different spheres of function (i.e. sensory, motor, pain and lifestyle) were found to be less well or poorly correlated. This finding is not novel, as several clinical and experimental nerve injury outcome studies have noted poor correlation between different outcome modalities. At best, sensory, motor, and pain/lifestyle measures correlated moderately with one another.

In a study of patients suffering from carpal tunnel syndrome, Levine et al examined the correlation between FSS, SSS, s2-PD, grip strength, key pinch strength and monofilament. They discovered that SSS and FSS correlated moderately with grip and key pinch strength. We have found modest correlation (r > 0.50, p ≤ 0.050) between these outcome measures as well within patients with ulnar neuropathy. Surprisingly (given the median nerve’s importance in subserving sensory hand function),
Levine found that FSS and SSS did not correlate with the two sensory modalities, s2-PD and monofilament testing. However, in a study performed by Chassard et al involving patients with median and ulnar nerve injuries, a strong relationship between functional status and two-point discrimination was observed, yet two-point discrimination was poorly reflective of outcome of similar patients in Rosen’s study. Our results indicate that both FSS and SSS correlate with s2-PD (r > 0.4, p < .025), and that SSS correlates moderately with monofilament (r = 0.42, p ≤ 0.05). These results are in agreement with Chassard’s and thus suggest that functional status and severity of symptoms (as assessed by the Levine method) do correlate with s2-PD in ulnar neuropathy patients. Further research is required to confirm Levine’s results and determine whether the FSS/s2-PD correlation is indeed absent in the patients with carpal tunnel syndrome.

Previous studies have not resolved whether s2-PD reflects patient outcome. For example, Foster and Edshage measured outcome in terms of pain, weakness and sensory deficits and reported that s2-PD had no association with outcome. Furthermore, Rosen et al have stated that s2-PD is not responsive to clinical change and have suggested that this tool be excluded from a nerve repair assessment battery. Conversely, Nathan et al found that normal s2-PD preoperatively predicted better long-term improvement (resolution of numbness or tingling (sensory or pain improvement)) using regression analysis and Greenwald et al have noted a statistically significant beneficial change in s2-PD following anterior submuscular transposition of the ulnar nerve. Our data indicate that s2-PD has moderate correlation between sensory, motor, and pain/lifestyle outcome measures.

Few studies have set out to determine which outcome measures are actually sensitive to clinical change, and those that have endeavoured to do so have likewise employed an inadequate number of assessment measures. Levine et al have determined that their symptom severity and functional status scores can be sensitive to clinical change, but they did not examine whether other measures might be sensitive as well. Rosen et al looked at the ability of a small number of tests to detect clinical change within patients with repaired ulnar and median nerves but, in initial studies, did not examine relationships between the various tests. In the recent past, Rosen and co-workers have examined these relationships and attempted to validate a model of functional outcome, by factor analysis. Interestingly (and similar to us), Rosen reports that three domains, sensory, motor, and pain/discomfort account for 73% of the variance.

Patient-specific index instruments have become popular to assess outcome. Several reports suggest that PSM are sensitive to clinical change, but these studies were not concerned with ulnar neuropathy. In our small cohort, PSM appear to reflect clinical status extremely well (Figure). We found that Lifestyle/Pain instruments were not only strongly correlated to each other in the first and in subsequent visits, but that they also appeared to reflect patient outcome accurately. In the present (limited) study, factor scores demonstrated a statistically detectable improvement over time, underscoring their likely strong sensitivity to change. Examining the sensitivity of the various outcome measures and the described factor scores in detecting clinical change in ulnar neuropathy should be further considered in future studies of ulnar neuropathy at the elbow.

The main limitation in this study was the small size of the patient cohort. Factor analysis requires a sample size of at least 150 cases to be conclusive. Acquiring such a large patient cohort would require a multi-institutional effort to concurrently administer the outcome measures and collect data. Once a large sample size is obtained, it will be possible to identify the unique combination of outcome measures which accurately depict clinical status and progress, and thereby develop a concise ulnar neuropathy battery. Ideally, this small battery of tests should be easy to administer, sensitive to clinical change and capable of capturing the overall trend in patients’ progression through an ulnar neuropathy. We hope that the identified battery of outcome measures will serve as the basis for a comparison of the superiority of different surgical interventions in the future.

ACKNOWLEDGEMENT

The authors thank Drs. Paul Binhammer, Brent Graham and Eugene Wai for providing critique and helpful suggestions to revise the manuscript and Marko Katic for assistance with statistical analysis.

REFERENCES