Author's response to reviews

Title: Diagnostic accuracy of the neurological upper limb examination I: Inter-rater reproducibility of selected findings and patterns.

Authors:

Jorgen R Jepsen (jri@ribeamt.dk)
Lise H Laursen (lisehl@fmm.sdu.dk)
Carl-Goran Hagert (carl-goran.hagert@ort.lu.se)
Svend Kreiner (s.kreiner@biostat.ku.dk)
Anders I Larsen (ail@novozymes.com)

Version: 2 Date: 15 December 2005

Author's response to reviews:

The following comments are referring to two submitted manuscripts:

* This manuscript on reliability (MS: 1303123447998486 - Diagnostic accuracy of the neurological upper limb examination I: Inter-rater reproducibility of selected findings and patterns)
* A manuscript on validity (MS: 1051549914800868 - Diagnostic accuracy of the neurological upper limb examination II: Relation to symptoms of patterns of findings).

Comments to review by Michele Sterling.

Discretionary revisions.
1. As requested by the Reviewer, further details have been supplied in the text of both manuscripts in terms of classification of responses to vibration and mechanosensitivity testing. Ratings from "no/normal" to "marked" were based on the response of the participants as well as interpretation of examiners.

Comments to review by John Quintner.

Major revisions.
1. Is the question posed by the authors new and well defined? Dr. Quintner summarizes the three principal questions on validity issues raised in the two manuscripts.
   a. The reproducibility of findings (i) has been dealt with in the reliability manuscript. I understand that the reviewer is satisfied with the text.
   b. Whether patterns of neurological dysfunction so identified are indicative of specific neurological diagnoses such as focal neuropathies (ii). This has been dealt with in the two primary manuscripts.
   c. Whether other diagnostic techniques such as MRI and electrophysiological studies can be utilised (iii). This issue has been dealt with in the primary version of the validity manuscript. I understand, however, that the Reviewer requests this question further elaborated to include a discussion with regard to indications for treatment based on this physical assessment. In the revised validity manuscript these aspects have now been included in the Discussion and summarized in the Conclusion.
2. Are the methods appropriate and well-described, and are sufficient details provided to replicate the work?
   a. We agree that it would be interesting for readers to get a more complete and detailed picture of the composition of patients in order to assess the potentials of the examination. Unfortunately, we did not register the requested additional data in our data base. Hoping that readers will get a better impression of the sample-composition we have, however, added the following text to the discussion in both manuscripts: "The symptomatic patients referred for assessment in occupational medicine did not merely represent a group of chronic pain patients. While some patients presented with long-lasting and major disabling symptoms others have had minor symptoms for a short period of time. The duration of upper limb symptoms ranged from a few months to several years preceding referral. About half of the patients were on sick-leave while the remaining patients were able to continue their work. Most patients with upper limb symptoms were formerly diagnosed with specific disorders such as tennis elbow or shoulder tendinitis. Many had several such diagnoses suggested by various specialists. Others were labelled as non-specific upper limb conditions such as RSI (repetition strain injury). In many patients a neuropathic condition was suspected and electrophysiological studies (mostly of the median nerve in the carpal tunnel) and imaging (especially of the cervical spine) performed. These additional diagnostic studies did not contribute
b. Dr. Quintner states correctly that grading of presence of changes in sensibility and nerve trunk
mechanosensitivity could clearly be simplified into Yes and No. As noticeable from the Tables 3 - 4 in the
reliability manuscript we have done both. In addition, we have presented dichotomized data for the patterns
(Table 5). The dichotomized data for sensibility and mechanosensitivity were employed for the kappa
analyses. However, for the study of correlation between patterns and symptoms the additional use of
graded data enables us to better reflect the inter-rater reliability without predefined cut-off values.

3. Is the data sound and well controlled?
a. Table 5 in the reliability manuscript and Table 2 in the validity manuscript describe the overlapping
neurological patterns found by the blinded examiners. The two examiners agreed on the presence of 90
patterns when all individual patterns are added. We understand the apparent discrepancy with comparison
to Table 4 in the validity paper in which the examiners agreed on the identification of the presence of any
pattern in only 30 limbs. The explanation is that in the 30 limbs a total of 90 locations of patterns were
identified. In the first version of the reliability manuscript this has been already explained. In order to avoid
misunderstandings we have explained findings further in the Result section in the validity manuscript.
b. In the Abstract to the validity manuscript the Result section states that the two examiners identified
pattern(s) suggesting focal neuropathy in 34/36 limbs, respectively, out of 38 symptomatic limbs. There is
actually no discrepancy to the 28 limbs agreed on according to Table 4. The 28 limbs constitute the
common part of the 34 and the 36 limbs, i.e. the number of symptomatic limbs with agreement between the
two examiners. We have explained further in the revised Abstract.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
a. Table 3 does in fact illustrate the relation of integrated findings to symptoms and to the diagnosis of the
primary examiners. It was not our intention to illustrate the number of limbs in which any pattern was
identified by the primary examiner, but we can see that this can be misunderstood. For that reason we have
clarified the legend to Table 3.
b. We have added in the validity manuscript a discussion about the inability of the primary examiners to
detect the neuropathic conditions indicated by the secondary examiners taking into account that the
detected patterns could occur in isolation as neuropathic conditions or they could be combined with upper
limb conditions diagnosed by the primary examiners.
c. The statement in the abstract that the improved diagnostic confidence indicates construct validity of the
physical examination is challenged by Dr. Quintner. We agree that we cannot claim that we have
established construct validity and it was not our intention to do so. By using the term "indicate", we intended
to say that this study is supportive of construct validity. However, to conclude that construct validity is
present would demand numerous studies each contributing in various ways. This is stated in the Discussion
and Conclusion as well as in the Abstract of the revised validity manuscript. It is acknowledged that in the
absence of an adequate "golden standard" such as is the case here, the question of validity is a real
challenge. While determination of criterion validity would require a golden standard, this type of validity
cannot be established. Construct validity is not defined unanimously by all researchers. We have used the
definition of J M Last in the Dictionary of Epidemiology [1]: "The extent to which the measurement
corresponds to theoretical concepts (constructs) concerning the phenomenon under study", i.e. the degree
to which inferences can legitimately be made from the operationalizations in the study (blinded physical
findings) to a theoretical construct which may reflect those operationalizations (symptoms). In face validity,
one looks at the operationalization and sees whether "on its face" it seems like a good translation of the
construct. We consider our contribution to be more than face validity. We have modified the text in order to
more precisely express that this study has contributed to the establishment of one aspect of construct
validity of our examination in terms of documenting the relation to the presence of symptoms. We hope this
is satisfactory for the Reviewer and the Editor.
d. For several reasons aiming to locate pathology is important. Knowledge about the location of pathology,
e.g., at the brachial plexus as suggested in many of the studied patients, may contribute to theories with
regard to pathophysiological mechanisms, and may eventually open new options with regard to primary,
secondary and tertiary prevention as well as to treatment. E.g. the current experience of the authors that
many of these patients seem to benefit from physiotherapy based on the adverse nervous tension concept
and actually focussing on mobilization of the brachial plexus is supportive of the presented findings but out
of the scope of this study. We acknowledge that treatment issues are fundamental for the clinical feasibility
and wish to continue research taking these aspects into consideration. In the two articles, however, we have
decided not to deal with treatment because this issue would rely on further validation of our physical
examination.
6. Do the title and abstract adequately convey what has been found?
a. Taking into consideration the potential influence of confounders, we have modified and rephrased the
text in the Discussion of the validity manuscript.
7. Is the writing acceptable?
a. We have made all the suggested minor changes with regard to the writing in the text of the present
manuscripts.

We are grateful for the valuable comments and suggestions provided by the reviewers and sincerely hope
that our submitted revisions shall turn out satisfactory to the Reviewers and the Editor.

Jorgen Riis Jepsen