Reliability, validity, and sensitivity to change of the Cochin hand functional disability scale in hand osteoarthritis


*The Réseau Rhumato Group
†Department of Physical and Rehabilitation Medicine, Hôpital Cochin, Université René Descartes, Paris, France
‡Department of Biostatistics, Hôpital Necker, Université René Descartes, Paris, France

Summary

Objective: To assess the reliability, validity and sensitivity to change of the Cochin hand functional scale in hand osteoarthritis (OA).

Background: The Cochin hand functional disability scale has been validated in rheumatoid arthritis.

Design: Patients with hand OA according to Altman’s criteria were included. Impairment outcome measures (VAS of pain, hand score of tenderness, clinical hand score of impairment, Kallman’s radiographic scale), functional disability measures [Cochin scale, Revel’s functional index (RFI), Dreiser’s functional index (DFI)] and patients’ perceived handicap (VAS) were recorded twice, at baseline and at a 6-month follow-up visit. Interobserver reliability was assessed using the intraclass correlation coefficient (ICC) and the Bland and Altman method. Construct (convergent and divergent) validity was investigated using the Spearman rank correlation coefficient and a factor analysis was performed. Sensitivity to change was assessed using the effect size (ES) and the standardized response mean (SRM), and the non-parametric Spearman rank correlation coefficient (r) was used to assess the correlation between quantitative variable changes and patient’s overall opinion.

Results: 89 patients (8 males, mean age 63 years) were included. Interobserver reliability was excellent (ICC=0.96). The Bland and Altman method showed no systematic trend. Correlations of the Cochin scale score with RFI (r=0.86), DFI (r=0.87), VAS of handicap (r=0.67), VAS of pain (r=0.54), tenderness (r=0.51), clinical impairment (r=0.32), and Kallman’s radiographic scale (r=0.13) indicated a good construct validity. Factor analysis extracted four main factors, accounting for 65% of the total variance. 51 patients were evaluated at the 6-month visit. The Cochin scale score had worsened with SRM and ES values of −0.26 and −0.17 respectively. Changes in the score had one of the highest correlation (r=0.47) with the patient’s overall opinion.

Conclusion: The Cochin hand functional disability scale which was first developed to assess the rheumatoid hand can be used to evaluate functional disability in hand OA. © 2001 OsteoArthritis Research Society International

Key words: Outcome measure, Disability, Osteoarthrosis, Hand.

Introduction

The hand is a common site of peripheral joint involvement in osteoarthritis (OA). The prevalence of OA of the hand has been estimated to be 38% and 24.5%, respectively, in women and in men over the age of 66 years. However, of those with radiological OA, symptomatic disease ranges from 20 to 40%. Although hand OA is often thought to make only a small contribution to overall disability, the handicap induced by limitations in performing activities of daily living such as dressing and feeding may be considerable. Few therapeutic (medical or surgical) trials on hand OA have been published, and even in most recent studies assessing surgical procedures or medical treatments disability outcome measures are not usually used. When disability is assessed, a generic instrument such as the health assessment questionnaire (HAQ) is used. The HAQ includes questions on hand activities but was not developed to assess hand OA. As there is no single functional status questionnaire to assess outcome in all situations of rheumatic disease, it seems preferable to use a specific instrument.

To our knowledge, only two questionnaires have been developed to assess functional disability in hand OA. One of them has never been published and its metric properties are uncertain. The other has been partly validated but its construct validity was not fully assessed and its sensitivity to change is unknown.

We have developed the Cochin scale, a practical functional disability scale for rheumatoid hands comprising 18 questions on daily activities. This scale is a hetero-questionnaire completed by the patient’s doctor. The principal qualities of this scale are its simplicity: it takes less than 3 min to complete and the total score is easily obtained by adding the score of each question rated on a Likert scale from 0 (done without difficulty) to 5 (impossible.
to do. Its reliability and validity have been demonstrated\textsuperscript{12} and its responsiveness has also been demonstrated in two situations, during the course of the disease\textsuperscript{13} and after hand surgery\textsuperscript{14}.

Several other hand disability scales have been proposed for rheumatoid arthritis (RA) patients\textsuperscript{15–20}. Most have not been properly validated and are not widely used or used only by their authors. To avoid increasing the number of outcome measures and to promote normalization of hand disability assessment in rheumatic diseases, it would be of interest to have a single instrument to assess hand disability in RA and in hand OA, the two most common rheumatic illnesses affecting the hands.

The aim of this study was thus to assess the inter-rater reliability, the construct validity and the sensitivity to change of the Cochin scale in hand OA.

Materials and methods

ELIGIBILITY CRITERIA

To be eligible for the study, the patients had to fulfill the criteria of hand OA defined by Altman\textsuperscript{21}. Patients were excluded on the basis of the following criteria: (1) severe psychiatric disorders (particularly psychosis and depression requiring a change in treatment in the last 30 days); (2) restricted hand motion due to skin lesions and Dupuytren’s contracture; (3) neurological disorders of the upper limbs; (4) upper limb arthroplasty, amputation, or joint fusion; hand and wrist surgery or trauma within 90 days of entry; and (5) inability to speak French fluently; (6) absence of hand X-ray in the last 6 months.

STUDY DESIGN

Outpatients were prospectively and consecutively recruited from one rehabilitation and five rheumatology departments over a 12-month period. They were evaluated for impairment, disability and perceived handicap by the physician to whom they had been referred (SP, XC, TC, RMF, FL, EN). A subgroup of patients underwent a second examination 6 months after the baseline visit. For the reliability study, performed in the rehabilitation department, the Cochin scale was administered twice at the baseline visit (SP, MMLC) at a 1 h interval. The scale was filled out at the beginning and at the end of the interview. In the interval, patients filled out several other questionnaires (with a total of 75 items) and had a physical examination.

The following parameters were recorded:

(1) At the baseline visit only, demographic and disease characteristics were recorded. Radiographic hand lesions were assessed using Kallman’s scale (range 0–198)\textsuperscript{22}.

(2) At both baseline and follow-up visits, disability and impairment measures and patients’ perceived handicap were recorded as described below.

Hand disability measures

The Cochin scale\textsuperscript{12} (appendix 1) is a heteroquestionnaire with 18 questions concerning daily living activities, each question scoring from 0, performed without difficulty, to 5, impossible to do. The total score is obtained by adding the scores of all questions (range 0–90).

The Revel functional index (RFI)\textsuperscript{23} is a heteroquestionnaire with 10 questions on daily activities each scoring from 0, performed without difficulty, to 2, impossible to do. The total score is obtained by adding the scores of all questions (range 0–20).

The Dreiser functional index (DFI)\textsuperscript{10} is an self-administered questionnaire with 10 questions on daily activities each scoring from 0, performed without difficulty to 3, impossible to do. The total score is obtained by adding the scores of all questions (range 0–30).

Impairment measures

Total score of tenderness as measured by the Ritchie articular index\textsuperscript{24} was recorded. Joints examined were the carpo-metacarpal and interphalangeal of the thumb and the proximal and distal interphalangeal of the fingers. The total score is obtained by adding the scores of each joint (range 0–60).

Clinical hand score of impairment was recorded. Joints examined were the carpo-metacarpal, metacarpo-phalangeal and interphalangeal of the thumb and the metacarpophalangeal, proximal and distal interphalangeal of the fingers. For each joint, the presence of nodes, swelling, deformation, and pain at passive motion was recorded and graded as 0, absence or questionable, 1, moderate, and 2, severe. The total score is obtained by adding the scores of each joint (range 0–240).

A visual analog scale was used to assess the intensity of pain in the hands (VAS pain)\textsuperscript{25}. It ranges from 0 mm (no pain) to 100 mm (maximum pain) and was recorded by the patients.

Patient’s perceived handicap

A visual analog scale was used to assess patient’s perceived handicap (VAS Hd)\textsuperscript{12}. It ranges from 0 mm (no handicap) to 100 mm (maximum handicap) and was recorded by the patients.

(1) At the follow-up visit only, the patient’s overall perception of hand disability compared with that at the baseline visit was recorded on an 8-level ordinal adjectival scale (disappeared, very much less severe, less severe, slightly less severe, unchanged, slightly more severe, more severe, very much more severe).

Statistical analysis

All analysis were performed using SAS software version 6.06 (SAS Institute Inc., SAS Campus Drive, Cary, NC 27513). Quantitative variables were described using mean±standard deviation (S.D.), minimum and maximum value. Qualitative variables were described using proportion and percentage.

Reliability

Interobserver reliability was studied using simultaneously the intraclass correlation coefficient (ICC)\textsuperscript{26} and the Bland and Altman method\textsuperscript{27} ; the two methods give complementary information as shown by Atkinson and I-Kuei Lin\textsuperscript{28,29}.
Construct validity

Construct validity was investigated in three ways. (1) Convergent validity was assessed by correlating the Cochin scale score with variables that should have a converging relationship. These variables were scores of the RFI and DFI, the VAS of handicap and the VAS of pain. (2) Divergent validity was assessed by correlating the Cochin scale score with variables known to have moderate or little correlation with disability. These variables were the scores of tenderness and clinical impairment, and radiological grading using Kallman’s scale. For convergent as well as for divergent validity, as a normal distribution could not be demonstrated for all the parameters studied, the non-parametric Spearman rank coefficient (r) was used to assess the correlation between two quantitative variables. Spearman’s coefficient values were interpreted as excellent >0.91, good 0.90–0.71, moderate 0.71–0.51, fair 0.50–0.31, or little or none 0.30. (3) Factor analysis was performed as follows: the principal component analysis was used to extract factors, the retained factors had eigenvalues>1. Then independent factors were obtained using the varimax rotation method.

Sensitivity to change

Five different statistical approaches were used to assess responsiveness31,32. Standardized response mean (SRM). The SRM is defined as the mean change in scores between the baseline and the follow-up visit divided by the standard deviation (S.D.) of the individual changes in scores. A higher SRM indicates greater responsiveness. A negative value indicates that the mean score at the baseline visit is smaller than the mean score at the follow-up visit.

Effect size (ES). The ES is defined as the mean change in scores between the baseline and the follow-up visit divided by the S.D. of the baseline score. A higher ES indicates greater responsiveness. A negative value indicates that the mean score at the baseline visit is smaller than the mean score at the follow-up visit.

Paired t-test was used to compare the mean of the individual changes of each subject to 0. The level of significance was chosen as α=0.05.

To assess the clinical relevance of the changes, Spearman’s correlation coefficient was used to study the relationship between the patient’s overall opinion index with the individual changes in each quantitative variable assessing disability, impairment and handicap.

With a responsive outcome measure, scores improve when the patient improves and worsen when the patient’s condition deteriorates33. SRM values were thus also calculated in the subgroup of patients who improved (patient’s overall opinion index answers at the follow-up visit: disappeared, very much less severe, less severe, slightly less severe) and in the subgroup of patients who deteriorated (patient’s overall opinion index answers at the follow-up visit: very much more severe, more severe, slightly more severe). The Mann–Whitney U test was used to compare changes in scores in the two groups of patients. The level of significance was chosen as α=0.05.

Finally, the Spearman rank correlation coefficient was also used to assess the pairwise correlation between individual changes in two quantitative variables, whether the variables assess disability, impairment or handicap.

Results

DEMOGRAPHIC AND CLINICAL DATA

Eighty-nine patients (81 women) were evaluated at the baseline visit. Their mean age was 63.2±8.9 years (range 44–81 years). Disease duration at the first visit was 9.9±9.4 years (range 0–41). Table I shows the statistical parameters (mean, standard deviation, minimum, maximum) of the scores of the disability measures (Cochin scale, the RFI and DFI), the impairment measures (VAS of pain, hand score of tenderness, clinical hand score of impairment, Kallman’s radiographic scale), and of handicap (VAS Hd).

RELIABILITY

The Cochin scale was administered twice at the baseline visit at a 1 h interval to the 41 patients recruited in the rehabilitation department by two of the authors (SP, MMLC). Inter-rater reliability was excellent with an ICC value of 0.96. The Bland and Altman method (Fig. 1) showed that the mean of the difference was 0.2 (S.D., 3.6). The distribution of the differences was homogeneous with no systematic trend (r=0.03). The limits of agreement for the two observers were 0.2 (mean)−7.1 (1.96 S.D.) =−6.9 and 0.2+7.1=7.3.

CONSTRUCT VALIDITY

Table II shows the results of the convergent and divergent validity. The scale had good convergent validity with

<table>
<thead>
<tr>
<th>Table I</th>
<th>Disability impairment and handicap scores of the 89 patients evaluated at the first visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Cochin scale score (range 0–90)</td>
<td>18.73</td>
</tr>
<tr>
<td>RFI score (range 0–20)</td>
<td>3.47</td>
</tr>
<tr>
<td>DFI score (range 0–30)</td>
<td>7.81</td>
</tr>
<tr>
<td>VAS pain (range 0–100)</td>
<td>44.60</td>
</tr>
<tr>
<td>Tenderness (range 0–60)</td>
<td>8.79</td>
</tr>
<tr>
<td>Clinical impairment (0–240)</td>
<td>27.26</td>
</tr>
<tr>
<td>Kallman’s index score (0–198)</td>
<td>70.95</td>
</tr>
<tr>
<td>VAS Hd (0–100)</td>
<td>43.77</td>
</tr>
</tbody>
</table>

S.D.: standard deviation; Min: minimum; Max: maximum; VAS: visual analog scale; DFI: Dreiser’s functional index; RFI: Revel’s functional index; Hd: handicap.
the other disability scales (RFI, DFI) and patient’s perceived handicap (VAS Hd) while its correlation with pain and impairment measures was only moderate (VAS Pain, tenderness), fair (score of clinical impairment), or none (Kallman’s radiographic score).

Factor analysis (Table III) extracted four factors with eigenvalues>1, which accounted for 65% of the total variance. The first factor represents activities requiring grip strength, the second, activities requiring dexterity and precision, the third, activities requiring pinch strength and the fourth, activities requiring pinch dexterity of the dominant hand. Table IV shows the loading of each question after varimax rotation on the four factors.

**SENSITIVITY TO CHANGE**

Fifty-one patients had a second evaluation during a follow-up visit at an interval of 6.3±0.5 months (range 5–7 months). Table V shows the statistical parameters (mean, standard deviation, minimum, maximum) of the scores of the disability measures (Cochin scale, the RFI and DFI), the impairment measures (VAS of pain, hand score of tenderness, clinical hand score of impairment, Kallman’s radiographic scale), and of handicap (VAS Hd) at the baseline and follow-up visits and the differences (mean, standard deviation, minimum, maximum, Student’s t-test P value) in scores between the first and the second visit.

For each patient, a negative value means that the score of the evaluated measure was higher at the second assessment. The Student’s t-test P value for mean differences of the Cochin scale score was 0.06, indicating that worsening of the disability nearly reached statistical significance. The responsiveness of the measures evaluated by the SRM and the ES is shown on Table VI. The values observed for the Cochin scale total score were among the highest of the parameters studied, with −0.26 for SRM and −0.17 for ES. The SRM and ES values of the Cochin scale score were higher than those observed for the two other disability scales (RFI, DFI) used in this study (−0.004 and −0.003, −0.03 and −0.02 for SRM and ES values of the RFI, and the DFI respectively), and those of the EVA of pain (−0.10 and −0.12 for RSM and ES respectively), and of the same order than those of the EVA of handicap (−0.25 and −0.28 for RSM and ES respectively).

Individual changes in the scale score had the second highest correlation with the patient’s overall assessment (r=0.47) (Table VII), suggesting that these changes were clinically relevant.

Table VIII summarizes the individual changes in scores in the group of patients who improved (22 patients) and
who deteriorated (19 patients). The Cochin scale discriminates well between those who improved and those who deteriorated (SRM = 0.30 and 0.75 respectively; Mann-Whitney test P value = 0.0002 between means of individual changes in the two groups).

Finally, individual changes in the scale score were best correlated with individual changes in the DFI, VAS Hd, and VAS of pain scores (r = 0.65, 0.59, and 0.57 respectively) (Table VII).

Discussion

This study demonstrates that the Cochin scale has excellent reliability, a good construct validity, and is a responsive outcome measure in hand OA.

Clinicians need to assess regional disability to measure patient limitations and clinical evolution before they can propose or evaluate local treatment. During the last decade, disability outcome measures have been progressively added to the evaluation of patients with rheumatic diseases especially in RA. However, assessment of hand disability using specific instruments is not usually included among the variables used for clinical evaluation in hand OA. Thus, disability induced by hand OA is not precisely determined. By contrast, it has been shown that hand involvement is an important component of disability in RA and that the score of the hand functional index (9 first questions of the Keitel functional index) explained over 60% of the variance of the whole index. The mean score of the Cochin scale in this study is equal to that observed in a previous study in RA patients. This result suggests that the disability caused by hand OA, in the population of patients assessed, is as severe as that observed in a population of RA patients with a mean disease duration of 10 years. This result must be confirmed by others but it highlights the need to assess disability in this condition.

It is unlikely that the high inter-rater reliability could be due to the short interval between the two tests. The Cochin scale was completed at the beginning and at the end of the interview. In the interval, patients filled out several other questionnaires (with a total of 75 items) and had a physical examination. Patients could remember some questions but were unlikely to remember previous answers. Moreover, the intra-class correlation coefficient value was the same than the one observed in a previous study where patients were interviewed at a 24-hour interval.

The validity of a questionnaire is mainly assessed by criterion referenced validity and by construct validity. As there is no gold standard to assess functional disability, we assessed construct validity. The Cochin scale has good correlation with RFI and DFI, two scales assessing functional disability, while it has moderate, fair, or little correlation with outcome measures assessing impairment, thus reflecting a good construct validity. These results confirm those of our previous studies in RA showing that disability scores or their changes over time are poorly correlated with impairment measures. The good correlation observed between the scale score and the VAS of handicap suggests that functional disability assessed by the Cochin scale reflects the patient’s perceived handicap.

Four factors were extracted by factor analysis. The first was activities requiring grip strength which seems to be the most severe type of disability for our patients. As shown in earlier studies, items are grouped according to the kind of movement required.

The factors extracted for patients with hand OA are not the same as those extracted in our previous study in RA. This is not surprising since joint damage differs in the both diseases, involving the wrist in RA but not in hand OA. In fact in RA, the first factor extracted was activities requiring stability and mobility of the wrist. This observation points out one advantage of this scale over global evaluation such as VAS of handicap. Our scale indicates the kind of activities that result in functional disability while VAS of handicap cannot provide detailed information.

Our third goal was to test the responsiveness of the Cochin scale during the course of the disease. Our hypothesis was that, without specific hand treatment such as local injections or surgery, only small changes would occur in our population during a 6-month period. Thus, the values of the SRM and the ES of all the evaluating measures tested in
this study are low since we tried to detect small changes. For the scale, SRM and ES values are of the same order as those observed over a 15-month period in RA patients. As several measures evaluated in this study such as pain and tenderness have SRM and ES values of the same magnitude as those of the scale score and as these measures have already been tested for their responsiveness and shown to have marked sensitivity to change after the introduction of a slow-acting antirheumatic drug or after surgery, one can assume that our scale has good sensitivity to change. Finally, it is unlikely that changes observed in the scale score are due to a lack of reliability as interobserver reliability is excellent.

As a responsive outcome measure should be able to discriminate between the subgroup of patients who improved and the subgroup of patients who deteriorated, we compared the SRM and the scores of the outcome measures in the two groups of patients. The Cochin scale was the outcome measure that best discriminated between patients who improved and deteriorated, confirming that it is a responsive instrument.

The study of the correlation between changes in disability and impairment scores revealed that they were poorly correlated. While clinical impairment and tenderness improved, functional disability measures deteriorated. These results confirm those of our previous studies showing poor or moderate correlation between disability scores and impairment measures in RA. They suggest that evaluation of treatment and therapeutic decisions in hand OA should not depend only on impairment measures.

The changes in the scale score have the best correlation coefficient with changes in patients’ perceived handicap. As the patient’s opinion of his disease can be considered an external standard to evaluate the relevance of the change, our results strongly suggest that the changes observed with the scale score are clinically meaningful.

In conclusion, the Cochin scale is a reliable, valid, and responsive instrument for the assessment of disability in hand OA. Functional disability can increase while impairment improves and thus should be systematically evaluated. This scale should help to assess the effectiveness of physical therapy, adaptive devices, local injections and hand surgery in terms of disability in hand OA. Finally, a single instrument can be used to assess disability in both RA and hand OA.

**Acknowledgments**

The authors wish to thank the members of the ‘Réseau Rhumatologie’ group for their advice and support; Laboratoires Searle for their financial support, and Catherine Payan for statistical assistance. Physicians of the ‘Réseau Rhumatologie’ are: Maxime Bréban, Alain Cantagrel, Anne Cotten, Philippe Goupille, Pascal Gugenbuhl, Erick Legrand, Virginie Legrange-Daumen, Laurence Lequen, Florence Levy-Well, Jean François Maillefert, Charles Masson, Rémy Nizard, Christian Pancaretti, Edouard Pertuiset, Xavier Puechal, Philippe Ravaud, Alain Sarad, Thierry Schaeverbeke, Jean Sibilia, Martin Soubrier and Eric Toussirot.
References

37. Kalla AA, Smith PR, Brown GMM, Meyers OL, Charlton D. Responsiveness of Keitel functional index compared with laboratory measures of disease activity in...

Appendix 1: The cochin scale

Answers to the questions:
0=Yes, without difficulty;
1=Yes, with a little difficulty;
2=Yes, with some difficulty;
3=Yes, with much difficulty;
4=Nearly impossible to do;
5=impossible to do.

In the kitchen
1. Can you hold a bowl?
2. Can you seize a full bottle and raise it?
3. Can you hold a plate full of food?
4. Can you pour liquid from a bottle into a glass?
5. Can you unscrew the lid from a jar opened before?
6. Can you cut meat with a knife?
7. Can you prick things well with a fork?
8. Can you peel fruit?

Dressing
9. Can you button your shirt?
10. Can you open and close a zipper?

Hygiene
11. Can you squeeze a new tube of toothpaste?
12. Can you hold a toothbrush efficiently?

At the office
13. Can you write a short sentence with an ordinary pen?
14. Can you write a letter with an ordinary pen?

Other
15. Can you turn a round door knob?
16. Can you cut a piece of paper with scissors?
17. Can you pick up coins from a table top?
18. Can you turn a key in a lock?