parkinsonian and tremulous. Her sleep pattern was disrupted. She was prescribed cholinesterase inhibitors, leading to some improvement. There was a previous medical history of an acute onset of aphasia and limb weakness, ascribed to a stroke. These symptoms had apparently resolved prior to the onset of her progressive illness. There was no family history of dementia or other neurodegenerative disease.

Patient 2 presented with gradual mental and physical slowing aged 69 years. His gait was shuffling, his hands trembled and his voice was hypophonic. Speech content was rambling and off the point. His personality was retained and he exhibited no stereotypic behaviours or altered eating habits. During a routine hospital admission he suffered an acute confusional episode lasting several days in which he became deluded and hallucinated. He believed he was being poisoned and that there were criminal activities taking place on the ward. He saw animals and puppets. There was no relevant previous medical or family history. Neurological examination revealed parkinsonism, but no worsening or fasciculations. Neuropsychological examination revealed performance variability, with fluctuations of attention and distractibility. Mild deficits were elicited in spelling, calculation, visual perception, spatial and constructional skills. Confrontation naming, word and sentence comprehension and repetition were preserved. He had a normal span of eight digits. There was no gestural apraxia. Memory recall was poor and on story recall he misconstrued the thematic content. However, he was well oriented, recognition memory was intact with no abnormal information loss over a delay. Executive tests elicited poor sequencing and a mild perseverative tendency, but preserved abstraction and mental set shifting. Category and letter fluency performance was within normal limits. Motor responses were slowed. A CT scan showed generalised cerebral atrophy. He was lost to follow-up.

**DISCUSSION**

Evidence was sparse for a role in DLB of repeat expansions in the C9ORF72 gene. One hundred patients showed no repeat expansion, despite a high prevalence of psychotic features. However, expansions were detected in two patients. The clinical picture in one was compounded by a previous vascular event. In the other, clinical features were not entirely typical. Psychotic features were transient. Performance fluctuations might feasibly have been products of distractibility. Neither patient had pathological confirmation of diagnosis. The question arises whether the clinical diagnosis of DLB was correct.

Neither patient had notable personality or behavioural change to suggest FTD. The posterior hemisphere imaging changes in one patient and generalised atrophy in the other would favour DLB rather than FTD. It is unclear whether C9ORF72 repeat expansions are fully penetrant or might be unrelated to the disease process. However, attention has previously been drawn to a phenotype of FTD that mimics DLB.5 The possibility exists that the two patients are further rare examples of TDP-43 proteinopathy, presenting with a DLB rather than FTD clinical phenotype.

In this study of DLB, visual hallucinations predictably predominated over delusions. By contrast, psychotic features associated with C9ORF72 repeat expansions, consist predominantly of delusions, particularly of a paranoid and somatoform type.4 This clinical distinction might suggest a different underlying neurobiological substrate.

In conclusion, C9ORF72 repeat expansions are not a major determinant of the psychotic features of DLB. However, the findings do not rule out the possibility that C9ORF72 repeat expansions can be associated with DLB. Alternatively, FTD may occasionally present as a more ‘posterior hemisphere’ disorder than previously thought, with clinical features resembling DLB. Repeat expansions in C9ORF72 may be a predictor of such atypical forms of FTD.

**REFERENCES**


How reliable is repeated testing for hemispatial neglect? Implications for clinical follow-up and treatment trials

**INTRODUCTION**

Patients with hemispatial neglect following right hemisphere brain damage fail to spontaneously orient towards or respond to...
contrallosal stimuli. The diagnosis and longitudinal assessment of the syndrome is not always straightforward. This is mainly due to two reasons: the heterogeneity of the syndrome and inter-individual differences in the time course of recovery from the disorder.

The neglect syndrome affects various cognitive components across patients, and one patient may show neglect on certain tasks but not on others. Because there is no single test able to detect neglect in all patients, a battery of several paper-and-pencil tests is usually required. However, little is known about their use as a tool for longitudinal assessments. This is of high clinical importance as repeat assessments are necessary to monitor changes in neglect severity related to spontaneous remission or a specific treatment. If a test per se is not 'stable', the variation in test results over repeated sessions may simply reflect low test–retest reliability and not the actual change of the underlying disorder.

We therefore investigated the test–retest reliability of three of the most commonly used paper-and-pencil tests for hemispatial neglect over several daily test repetitions in chronic neglect patients.

METHODS

Subjects

Fifteen patients (mean age 57.4 years, SD 14.2) with left hemispatial neglect following right hemisphere stroke (median time 7.6 months since stroke) participated in this study (see online supplementary table 1 and online supplementary figure 1 for details).

Hemispatial neglect was diagnosed if patients showed at least two more omissions on the left than on the right in standard cancellation tasks (Bells and Mesulam).

Patients provided written informed consent before participating in this study which was approved by the National Research Ethics Service.

Testing procedure

Each patient was tested daily on five consecutive days, always at the same time of the day, on the following three paper-and-pencil tests.

Line bisection

Three horizontal black lines of 17 cm length should be bisected, each being presented on a separate white sheet of paper (A4 size). Total deviation from the actual centre of the lines (mm) was measured and a mean was calculated for each session.

Bells test

Patients should circle bells (n=35) among 280 distractors presented on an A3 sized sheet of paper. Again, the total number of cancellations and the lateralisation index were calculated for each session.

Mesulam and Weintraub's symbol cancellation test

Patients were asked to circle targets (n=60, >300 distractors) on an A3 sized sheet of paper. The ICC was high for total number of cancellations and calculated a spatial lateralisation index: number of left-sided subtracted from the number of right-sided cancellations divided by total number of cancellations.

RESULTS

To assess test–retest reliability, traditional correlation coefficients such as Pearson's r are considered inappropriate as they ignore the degree of absolute agreement between repeated testing. Using SPSS V15.0 we calculated the intraclass correlation coefficient (ICC), a measure that accounts for the relative and absolute reliability. The ICC produces a reliability index between 0 and 1, with <0.7 indicating poor and >0.7 good reliability.

The ICC was high for 'total number of cancellations' (0.84 and 0.83 for Bells and Mesulam, respectively) and 'lateralisation index' (0.83 and 0.82), indicating high test–retest reliability for the cancellation tasks (table 1, online supplementary figure 2).

In contrast, ICC was far lower for the line bisection task (0.47). It also did not increase when the analysis was based on dichotomous outcomes (normal vs pathological 'neglect-like' deviation, with a cut-off point of +5.5 mm, ICC=0.41, 95% CI 0.19 to 0.69).

As shown in table 1, we also analysed the SEM and the 95% CI for test–retest agreement (CI, twice the SEM). These useful 'score bands' give an absolute range of test scores in which the true score of a subject on retesting will most likely fall into.

SEM was calculated as SEM = s × \sqrt{1 - ICC}, where s is the SD for the test, derived from patients' performances on all sessions.

DISCUSSION

In this study, we investigated the test–retest reliability of three major neglect tests. In contrast to previous studies with only one test repetition following a brief time interval, we chose a more clinically realistic scenario with repeated assessments over consecutive days. We examined only chronic neglect patients; thus, variability in test performance could be related to the characteristics of the measurement as opposed to fluctuations of the underlying disorder.

We found that both cancellation tasks (Bells and Mesulam) were very stable on retesting. In contrast, neglect patients' performance on line bisection fluctuated considerably over sessions, demonstrating low test–retest reliability.

Our study further provides 'error bands' in absolute test units that are clinically more useful than reliability coefficients. In the Bells test, for instance, a difference in test performance between test and retest of more than five cancellations is most likely due to a change in neglect severity and not test variability. In contrast, a reduction of about 15 mm deviation in line bisection does not necessarily mean neglect improvement but can be due to test variability.

Notably, test–retest reliability of the line bisection task also did not increase when the analysis was based on a cruder dichotomous score (instead of absolute millimetres), ruling out different granularities of the measures as an explanation. Instead, beside learning and fatigue effects, one reason for the low test–retest reliability of the line bisection task may be its lower sensitivity and specificity for hemispatial neglect when compared with cancellation tasks. Line bisection can also be influenced by other conditions sometimes associated with spatial neglect (eg, visual field defects). It may explain the double dissociation among individual patients with normal performance on one and pathological performance on the other test.

Regardless of the cause, a neglect test like the line bisection task with low test–retest reliability is certainly not a good tool for the longitudinal assessment of a patient's neglect severity in treatment trials or during

---

**Table 1** Mean scores, test–retest reliability and absolute agreement for the different tests in 15 neglect patients over five sessions

<table>
<thead>
<tr>
<th>Tests/parameters (score range)</th>
<th>Mean score (SD)</th>
<th>Reliability coefficient (ICC 95% CI, p&lt;0.01)</th>
<th>95% CI for test–retest agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bells test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of cancellations</td>
<td>18.3 (6.1)</td>
<td>0.84 (0.71 to 0.93)</td>
<td>2.43 ± 5</td>
</tr>
<tr>
<td>Lateralisation index (−1 to +1)</td>
<td>+0.49 (0.41)</td>
<td>0.83 (0.68 to 0.93)</td>
<td>0.17 ± 0.34</td>
</tr>
<tr>
<td>Mesulam cancellation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of cancellations</td>
<td>25.6 (12.4)</td>
<td>0.83 (0.68 to 0.93)</td>
<td>5.11 ± 10</td>
</tr>
<tr>
<td>Lateralisation index (−1 to +1)</td>
<td>+0.66 (0.39)</td>
<td>0.82 (0.67 to 0.93)</td>
<td>0.16 ± 0.32</td>
</tr>
<tr>
<td>Line bisection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deviation from the centre (−85 to 85 mm)</td>
<td>+4.3 (10.7)</td>
<td>0.47 (0.23 to 0.73)</td>
<td>7.81 ± 16 mm</td>
</tr>
</tbody>
</table>

ICC, intraclass correlation coefficient.
rehabilitation. Established cancellation tasks, in contrast, can be recommended as very reliable measures for the longitudinal assessment of neglect.

Björn Machner,1,2 Yee-Haur Mah,1 Nikos Gorgoraptis,1 Masud Husain1

1National Hospital for Neurology and Neurosurgery, UCL Institute of Cognitive Neuroscience and Institute of Neurology, Queen Square, London, UK; 2Department of Neurology, University of Lübeck, Lübeck, Germany

Correspondence to Dr Björn Machner, Department of Neurology, University of Lübeck, Ratzeburger Allee 160, D-23538 Lübeck, Germany; bjoern.machner@neuro.uni-luebeck.de

Additional materials are published online only. To view this file please visit the journal online (http://dx.doi.org/10.1136/jnnp-2012-303296).

Contributors BM, YM, NG and MH were all involved in the conception, design, acquisition, analysis and interpretation of data. BM wrote the first draft that was revised critically and finally approved by YM, NG and MH.

Funding This work was supported by The Wellcome Trust. BM was supported by a fellowship of the European Neurological Society (ENS).

Competing interests None.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the National Research Ethics Service (NRES).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The authors are open to discussion regarding data sharing.

Received 23 May 2012
Revised 19 July 2012
Accepted 1 August 2012

REFERENCES

How reliable is repeated testing for hemispatial neglect? Implications for clinical follow-up and treatment trials

Björn Machner, Yee-Haur Mah, Nikos Gorgoraptis, et al.

*J Neurol Neurosurg Psychiatry* 2012 83: 1032-1034 originally published online August 25, 2012
doi: 10.1136/jnnp-2012-303296

Updated information and services can be found at:
http://jnnp.bmj.com/content/83/10/1032.full.html

These include:

**Data Supplement**

"Supplementary Data"
http://jnnp.bmj.com/content/suppl/2012/08/24/jnnp-2012-303296.DC1.html

**References**

This article cites 5 articles, 2 of which can be accessed free at:
http://jnnp.bmj.com/content/83/10/1032.full.html#ref-list-1

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/