A new bedside test of gestures in stroke: the apraxia screen of TULIA (AST)

T Vanbellingen,1 B Kersten,2 A Van de Winckel,3 M Bellion,1 F Baronti,1 R Müri,4 S Bohlhalter4

ABSTRACT

Background Apraxia in patients with stroke may be overlooked, as clumsiness and deficient gestural communication are often attributed to frequently coexisting sensorimotor deficits and aphasia. Early and reliable detection of apraxia by a bedside test is relevant for functional outcome in patients with stroke. The present study was aimed at constructing a new bedside screening test for apraxia, called the Apraxia Screen of TULIA (AST), based on the comprehensive standardised Test for Upper-Limb Apraxia (TULIA).

Methods First, an item-reduction analysis of the TULIA (48 gestures) was performed, based on the methods of classical test theory and on a larger sample of patients with stroke (n=133) and matched healthy controls (n=50). Stepwise elimination of items resulted in a set of 12 items, demonstrating high internal consistency (Cronbach alpha=0.92). The six-point scoring method of the TULIA was dichotomised to the score levels pass and fail. In the second part of this study the validity of the AST was assessed prospectively in a new cohort of patients with stroke (n=31) by using the Pearson correlation analysis and binary classification display with the TULIA.

Results and discussion Validation of the 12-item AST with the TULIA showed a remarkable diagnostic reliability with high specificity, sensitivity and positive predictive value, for the presence and severity of apraxia. The AST is shown to be a reliable and valid bedside test in patients with stroke, allowing a straightforward assessment of apraxia within a few minutes.

SUBJECTS AND METHODS

Development of AST

The AST was developed in two phases. In a first phase, an item reduction in the TULIA was performed, based on the original sample (n=183); in the second phase, the validity of the AST was assessed, based on a new sample (n=31). The TULIA consists of 48 items, with an equal number of items testing for imitation and pantomime. In both of these task domains, the respective set of 24 items comprises eight meaningless gestures, eight intransitive (communicative) and eight transitive (tool-related) gestures. Based on separate cut-off scores and a detailed six-point scoring method (see online appendix 1), the TULIA can identify the pattern as well as the severity of an apraxia impairment. Statistical analyses were performed using PASW for Windows (Version 18.0.0; SPSS).

The item-reduction analysis was based on classical test theory. Accordingly, items with less inter-rater reliability, floor and ceiling effects, less internal consistency and lower content validity (based on expert opinion: TV and SB) were excluded stepwise.

The six-point scoring system of the TULIA was simplified to a dichotomous pass or fail as follows. Score levels 5 (normal movement), 4 and 3 of the TULIA were combined in AST to the score 1 (pass) (see online appendix 1). The decision to consider the TULIA score levels 4 and 3 as ‘normal
movement’ was based on the fact that these error types may be difficult to discern by non-specialist examiners. They either did not affect trajectory of gestures, that is, overall movement shape (including the plane relative to a tool or own body) or were subtle and corrected during ongoing movement. However, if errors affecting trajectory persisted, even though being only mild (score=2), gesture performance was considered deviant. Thus, the score levels 2, 1 and 0 of TULIA were pooled to the score 0 (=fail), a score that includes major spatial or semantic errors. Examples of apraxic errors observed in patients with stroke have been replicated by a model for different items and score levels; the videos are available at the author’s website (http://www.tulia.ch).

Finally, the cut-off levels of AST were determined in the original sample (n=133) by comparing the corresponding 12 items from TULIA recoded to pass and fail with the full version. Accordingly, using cut-off levels of 9 and 5, high specificity (93%) and sensitivity (88%) for mild and severe apraxia could be estimated.

For the validation of the AST test, patients with stroke were examined with both the AST and the TULIA. The diagnostic reliability of the AST was calculated using the Pearson correlation analysis, with the TULIA score and the AST score as dependent variables. Furthermore, a binary classification display of the AST and TULIA was used to demonstrate sensitivity, specificity and predictive values of the AST.

### Results

#### Subjects

The item-reduction analysis performed in phase I was based on the data we obtained from a sample of 133 patients (84 with a left-hemispheric stroke) and 50 healthy controls. Details of their clinical characteristics were reported elsewhere. Demographic and clinical data of the new patient controls. Details of their clinical characteristics were reported elsewhere.1 Demographic and clinical data of the new patient sample participating in phase II are summarised in table 1. The patients were selected prospectively, the selection criterion being that they had sustained a single radiologically confirmed stroke (CT or MRI) and did not suffer from any brain disorder; in particular, showed no signs of dementia (according to DSM IV criteria) or Parkinsonism, or suffered from orthopaedic disorders of the upper non-hemiplegic limb that could interfere with the TULIA or AST task. Furthermore, patients with impaired consciousness or an inability to understand basic task instructions were also excluded. All patients in phase II were right-handed, as assessed by the Edinburgh Handedness Inventory.17 Written informed consent for participation in the study was obtained from all patients according to the Declaration of Helsinki, 1975. The study was approved by the local ethical committee.

#### Procedure

All patients were examined by two investigators (TV and SB) with the TULIA test1 for apraxia, the Bell test18 for neglect and the 10 item screening version of the TOKEN test19 for aphasia. Performances of the TULIA were videotaped for later scoring. Within 24 h of examining the patients with these tests, the AST was administered and immediately scored by the same investigators. Scoring of the TULIA was done by experienced raters (FA and CH; see acknowledgements) who were blinded for the results of the AST.

To perform both the TULIA and the AST test subjects were either seated in front of the examiner or sat in an upright position in bed. In both of these tests the patients executed the movements with their non-paretic upper limb. The imitation tasks required the patients to reproduce the movement that the examiner had just demonstrated in a mirrored fashion. In the pantomime tasks the patients were given oral instructions as to the gesture they were to perform (see online appendix 2). The administration of the AST required on average 5 min after the patient was properly seated and had received the necessary instructions. The order of items followed the list in online appendix 2.

#### Results

The results of the stepwise item-reduction analysis are summarised in a flow chart (figure 1). Initially, four of the 48 items of the TULIA with lower inter-rater reliability (weighted kappa <0.60) were eliminated. Second, 12 items with a ceiling effect (mean score >4) were excluded. Third, an iterative elimination of 15 items with lower corrected item-to-total correlation coefficients (r<0.60) was carried out. The remaining 17 items had a high reliability coefficient (Cronbach alpha=0.91) for the temporary total scale. Fourth, the five items with lower content validity were eliminated. We were thus left with the 12 items that constitute the AST, seven items in the imitation and five in the pantomime domain. The item-reduction analysis resulted in the elimination of 15 meaningless, 13 intransitive and eight transitive gestures from the set of 48 items that constitute the TULIA. Thus, the AST consists of one meaningless gesture, three intransitive and eight transitive gestures. Internal consistency for the new scale AST (k=12) remained high (Cronbach alpha=0.92).

According to the dichotomous scoring system (1=pass, 0=fail) the maximum score for AST was 12. Furthermore, the cut-off scores were set at 9 and 5 for mild and severe apraxia, respectively (for rationale, see Subjects and methods section). On the basis of AST test scores, 19 patients (61%) were classified apraxic, and 12 patients (39%) non-apraxic. In the group with apraxia, 12 patients (59%) were mildly, and seven patients (25%) severely apraxic. All patients with severe apraxia had sustained a left-hemispheric stroke, whereas the group of mild aphasia also included two patients with a right-hemispheric stroke. As summarised in table 2, mean total scores and domain subscores of AST and TULIA were significantly lower in apraxia than in non-apraxic patients.

Further validation demonstrated a highly significant correlation between AST and TULIA scores (Pearson correlation

---

**Table 1** Characteristics of patients (phase II)

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>63.4 (13.7)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>21</td>
</tr>
<tr>
<td>Women</td>
<td>10</td>
</tr>
<tr>
<td>Type of stroke</td>
<td></td>
</tr>
<tr>
<td>Infarction</td>
<td>26</td>
</tr>
<tr>
<td>Haemorrhage</td>
<td>5</td>
</tr>
<tr>
<td>Hemisphere of stroke</td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>5</td>
</tr>
<tr>
<td>Left</td>
<td>26</td>
</tr>
<tr>
<td>Phase of stroke</td>
<td></td>
</tr>
<tr>
<td>Acute (&lt;4 weeks)</td>
<td>15</td>
</tr>
<tr>
<td>Chronic (&gt;4 weeks)</td>
<td>16</td>
</tr>
<tr>
<td>Aphasia (token test, screening version)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>8</td>
</tr>
<tr>
<td>Mild</td>
<td>4</td>
</tr>
<tr>
<td>Duration (poststroke to initial testing; days)</td>
<td></td>
</tr>
<tr>
<td>Acute</td>
<td>19.0 (5.3)*</td>
</tr>
<tr>
<td>Chronic</td>
<td>57.3 (25.5)*</td>
</tr>
<tr>
<td>Total</td>
<td>36.9 (25.7)*</td>
</tr>
</tbody>
</table>

*Mean (SD).
The present study was aimed at constructing a validated bedside test for apraxia in patients with stroke, referred to as the AST. The construction of the AST was based on an item-reduction analysis of a comprehensive test for gestures, called TULIA. For this apraxia test, we recently provided a clinimetric evaluation in a large cohort of patients with stroke and age-matched healthy subjects. The AST allows for a rapid assessment of apraxia, as the number of items, reduced from 48 to 12, is considerably lower than in the TULIA. Furthermore, the original six-point scoring method was dichotomised to fail and pass so that an ‘online’ judgement has become possible. The AST can be administered without video recordings which are required for differential diagnosis display (table 3).

**DISCUSSION**

As may be expected, within the group of left-hemispheric patients with stroke (n=26) the aphasics (n=12) had significantly lower scores in the AST than the non-aphasics (2.9±1.9 and 9.7±0.6, unpaired t test, p<0.001). To assess the potential confound of apraxia by aphasia or neglect, Pearson correlations were calculated between the AST, the Token test and the Bell test in the apraxia subpopulation (n=19). Total AST scores correlated significantly with the Token test (r=0.81, p<0.001). Further analysis showed that the association was particularly strong for the non-verbal imitation domain (r=0.89, p<0.001) and did not reach significance for verbally mediated pantomime domain (r=0.42, p=0.075). There was no relationship between the Bell test scores and the AST total scores and subscores (all p>0.05).

**Table 2** Mean (SD) scores of apraxic (n=19) and non-apraxic (n=12) group

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Imitation</th>
<th>Pantomime</th>
<th>Total</th>
<th>TULIA</th>
<th>Imitation</th>
<th>Pantomime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apraxic</td>
<td>4.3 (2.9)</td>
<td>1.5 (1.8)</td>
<td>2.5 (1.6)</td>
<td>126.3 (53.5)</td>
<td>72.8 (26.7)*</td>
<td>53.5 (31.2)*</td>
<td></td>
</tr>
<tr>
<td>Non-apraxic</td>
<td>10.4 (1.0)</td>
<td>6.3 (0.9)*</td>
<td>4.2 (0.8)*</td>
<td>210.6 (9.2)</td>
<td>106.3 (5.3)</td>
<td>104.3 (7.2)</td>
<td></td>
</tr>
</tbody>
</table>

Factorial mixed design ANOVA demonstrated a significant interaction effect between group (apraxic vs non-apraxic) and test (Test for Upper-Limb Apraxia (TULIA) vs Apraxia Screen of TULIA (AST)) (F(1,29)=24.58, p<0.001) for total scores. A significant interaction effect was also found between test (TULIA vs AST) and domain (imitation vs pantomime) (F(1,29)=6.47, p=0.02). Post hoc paired t tests revealed significant differences (p<0.01) between imitation and pantomime subscores of the AST for the non-apraxic group and of the TULIA for the apraxic group. There was no significant age difference between the apraxic and non-apraxic group (64.1 (16.2) and 62.3 (8.8), p=0.05). Maximum score of AST=12, maximum score of TULIA=240.
scoring schemes but not easily accommodated to bedside testing. The validation with the TULIA revealed the AST to be highly specific and sensitive. The AST proved to be of diagnostic relevance not only in assessing the presence of apraxia reliably within a few minutes, but also in providing a valid classification of the severity of the apraxia, as shown by the almost perfect categorical match between the AST and the TULIA scores. The lower negative predictive value suggests that mild forms of apraxia may be missed with the AST, probably due to the lower sensitivity of the dichotomous scoring method. Likewise, lower sensitivity of the AST may also account for the significant group and test interaction for total scores (table 2). Finally, the high correlation between the AST and the corresponding recoded items from TULIA were indicative of a high test—retest reliability.

Most patients with a left-hemispheric stroke and apraxia are aphasic. The strong correlation of the total AST scores with the Token test was not surprising. However, the presence of an aphasia is unlikely to account for the diagnosis of apraxia in our patient cohort, as a strong association was found between the total AST score and the Token test for the imitation part, which is basically independent of language processing. Nonetheless, aphasia with severe comprehension deficits may lead to confounds in praxis testing when it is a matter of performing pantomimes on verbal command. Thus, it is conceivable that severely aphasic patients may fall below the cut-off score for apraxia in the AST due to impaired auditory comprehension in the pantomime part. In selected cases, additional testing of the TULIA with separate cut-off scores for imitation may be needed to confirm the diagnosis of apraxia.

A limitation of the study is the possible bias by the examiner’s global impression from TULIA on his subsequent scoring of the AST, although the high correlation at the item level would not have been expected based merely on a vague impression. Another concern is the small sample size used for validation; it have been expected based merely on a vague impression. One concern is the possible bias by the examiner’s global impression from TULIA on his subsequent scoring of the AST, although the high correlation at the item level would not have been expected based merely on a vague impression. Another concern is the small sample size used for validation; it have been expected based merely on a vague impression. Nonetheless, aphasia with severe comprehension deficits may lead to confounds in praxis testing when it is a matter of performing pantomimes on verbal command. Thus, it is conceivable that severely aphasic patients may fall below the cut-off score for apraxia in the AST due to impaired auditory comprehension in the pantomime part. In selected cases, additional testing of the TULIA with separate cut-off scores for imitation may be needed to confirm the diagnosis of apraxia.

In conclusion, apraxia may go undetected in acute stroke or early rehabilitation units when sensorimotor and language impairments apparently predominate. AST offers a reliable and valid bedside instrument to accurately and quickly diagnose apraxia, which is important for planning targeted and comprehensive rehabilitation programmes early.

Acknowledgements We thank F Antonietti and C Held, for their crucial assistance in evaluating the videotapes. We are also grateful to J Milburn and D Weniger, for editing the manuscript, and H Mattie, for his advice.

Competing interests None.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the University of Leuven and University of Bern.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

A new bedside test of gestures in stroke: the apraxia screen of TULIA (AST)

T Vanbellingen, B Kersten, A Van de Winckel, et al.

J Neurol Neurosurg Psychiatry 2011 82: 389-392 originally published online October 9, 2010
doi: 10.1136/jnnp.2010.213371

Updated information and services can be found at:
http://jnnp.bmj.com/content/82/4/389.full.html

These include:

Data Supplement
"Web Only Data"
http://jnnp.bmj.com/content/suppl/2011/03/31/jnnp.2010.213371.DC1.html

References
This article cites 18 articles, 6 of which can be accessed free at:
http://jnnp.bmj.com/content/82/4/389.full.html#ref-list-1

Article cited in:
http://jnnp.bmj.com/content/82/4/389.full.html#related-urls

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.

Topic Collections
Articles on similar topics can be found in the following collections
Stroke (1303 articles)

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/