CLINICAL AND PROGNOSTIC PROPERTIES OF STANDARDIZED AND FUNCTIONAL APHASIA ASSESSMENTS

Ann Charlotte Laska, MD¹, Aniko Bartfai, PhD², Anders Hellblom³, Veronica Murray, MD, PhD¹ and Thomas Kahan, MD, PhD⁴

From the ¹Division of Internal Medicine and ²Division of Rehabilitation Medicine, Karolinska Institutet, Department of Clinical Sciences, Danderyd Hospital, ³Department of Speech Pathology, Danderyd University Hospital and ⁴Division of Cardiology, Karolinska Institutet, Department of Clinical Sciences, Danderyd Hospital, Stockholm

OBJECTIVE: To compare standardized and functional aphasia tests in patients after acute stroke.

DESIGN: Data were collected at baseline and at 6 months in 2 prospective single-centre studies: one observational study (study I, n = 119) and one randomized trial of moclobemide vs placebo (study II, n = 89).

SUBJECTS: Patients with aphasia after acute stroke.

METHODS: Degree of aphasia was examined using the Coefficient (Coeff) in Norsk Grunntest for Afasi (standardized) and the Amsterdam-Nijmegen Everyday Language Test (ANELT) (functional). Statistical comparisons were made using one-way analysis of variance and multivariate regression analyses.

RESULTS: The degree of aphasia measured with Coeff and ANELT correlated closely throughout the study (r² = 0.71–0.87, p < 0.0001). In study I, 24 patients recovered completely within 6 months. A Coeff ≥ 49 and ANELT ≥ 3.5 predicted complete recovery equally well. Coeff was sensitive to differentiate between patients with low values on ANELT, whereas ANELT was sensitive to differentiate between patients with high Coeff values.

CONCLUSION: The 2 tests show a close and consistent correlation over time and are equally sensitive to improvement. They have a similar capacity to predict complete recovery. A standardized test appears to be more suitable for patients with aphasia in the acute stage, while a functional test is more suitable in the subacute/chronic stage.

Key words: aphasia, stroke, standardized test, functional test, outcome.


Correspondence address: Ann Charlotte Laska, MD, Karolinska Institutet, Department of Clinical Sciences, Danderyd Hospital, Division of Internal Medicine, SE-182 88 Stockholm, Sweden. E-mail: ann-charlotte.laska@ds.se

Submitted July 25, 2006; accepted February 5, 2007

INTRODUCTION

Measurement of the degree of aphasia in acute aphasic stroke patients is necessary to obtain a baseline for evaluating outcome of treatment and is a significant subject characteristic in predicting prognosis. The degree of aphasia generally can be assessed by 3 groups of tests. The first group comprises the standardized aphasia tests, which measure language impairment. All standardized aphasia tests assess speech fluency, comprehension, naming and repetition, and some also assess reading and writing. The sum of the main parts is used as a measure of the degree of aphasia. Standardized aphasia tests also classify the type of aphasia, although the tests have different methods of determining the type of aphasia. According to some studies (1) standardized aphasia tests are less suitable for assessing functional recovery. Secondly, the degree of aphasia can be measured by functional communication assessments, which can be performed by speech therapists or significant others. Functional communication may be assessed verbally or non-verbally, but none of the functional tests are entirely non-verbal. The third group comprises neurological impairment tests, which usually have one or a few items on language impairment. Table I shows the most common standardized test for aphasia syndrome, functional communication tests, and the aphasia item in two neurological impairment tests. Considerable spontaneous recovery occurs in aphasia after stroke, especially over the first few months. The most important determinant for recovery is the initial severity of aphasia, but recovery is also influenced by age and whether the aphasia is fluent or non-fluent (2–4). Patients with milder aphasia have a greater chance of complete recovery. However, no study has shown how well tests can predict complete recovery. Thus, we studied patients with aphasia after acute stroke, and compared the Norsk Grunntest for Afasi (NGTA) (5) and the Amsterdam-Nijmegen Everyday Language Test (ANELT) (6).

METHODS

This report is based on the results of 2 prospective studies in patients with aphasia after acute stroke. Study I is an observational study in which consecutive patients were included during a 1-year period (4). Study II is a randomized, placebo-controlled, double-blind trial to evaluate the possible effects of moclobemide on recovery in the degree of aphasia (7). Hence, the sample size was determined by the primary aim of that study, as described elsewhere (7). A flow chart for both studies is shown in Fig. 1.

SUBJECTS

Study I included 119 patients with aphasia after acute stroke. They were evaluated with aphasia tests at baseline and at 3 and 6 months. The acute test was performed within a median time of 5 days (range 0–30 days) after stroke onset; 90% of patients were tested within 11 days.
Table I. Some tests for aphasia.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Reference</th>
<th>Time to perform (min)</th>
<th>Verbal / non-verbal</th>
<th>Classification of type of aphasia</th>
<th>Measures degree of aphasia</th>
<th>Tester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PICA</td>
<td>Porch 1971 (21)</td>
<td>60</td>
<td>Both / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>WAB</td>
<td>Kertesz 1982 (9)</td>
<td>60</td>
<td>Verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>BDAE</td>
<td>Goodglass &amp; Kaplan 1983 (22)</td>
<td>60</td>
<td>Verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>AAT</td>
<td>Huber et al. 1984 (23)</td>
<td>30–45</td>
<td>Verbal / non-verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>NGTA</td>
<td>Reinvang 1985 (5)</td>
<td>10–15</td>
<td>Verbal / non-verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>Functional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCP</td>
<td>Sarno 1969 (24)</td>
<td>45</td>
<td>Verbal / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>The speech questionnaire</td>
<td>Lincoln 1982 (25)</td>
<td>Not applicable</td>
<td>Both / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Significant others</td>
</tr>
<tr>
<td>CETI</td>
<td>Lomas et al. 1989 (14)</td>
<td>Not applicable</td>
<td>Both / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Significant others</td>
</tr>
<tr>
<td>ANELT</td>
<td>Blomert et al. 1994 (6)</td>
<td>10–25</td>
<td>Verbal / non-verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>ASHA FACS</td>
<td>Frattali et al. 1995 (26)</td>
<td>20</td>
<td>Verbal / non-verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>CADL 2</td>
<td>Holland 1980 (27)</td>
<td>25</td>
<td>Verbal / non-verbal</td>
<td>Yes</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>Neurological impairment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIHSS</td>
<td>Brott et al. 1989 (28)</td>
<td>5</td>
<td>Verbal / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Professional</td>
</tr>
<tr>
<td>SSSS</td>
<td>Röden-Jüllig et al. 1994 (11)</td>
<td>5</td>
<td>Verbal / non-verbal</td>
<td>No</td>
<td>Yes</td>
<td>Professional</td>
</tr>
</tbody>
</table>


**Fig. 1. Study I and II. In study I the final follow-up at 18 months has been omitted.**

The results of a final follow-up at 18 months have been reported elsewhere (4). Of the 119 patients included in the study, 24 recovered completely, 66 still had aphasia at 6 months, 13 were dead, 4 developed dementia and 12 were lost to follow-up at 6 months.

Study II included 89 patients with aphasia after acute stroke. They were tested for aphasia at baseline and at 6 months after stroke onset. For inclusion, the patients were required to have a degree of aphasia of 1.0–4.0 according to ANELT. The baseline test in study II was performed within a median time of 18 days (range 6–34 days) and 90% of the subjects were studied within 25 days. At 6 months, 76 patients could be tested, and 13 were lost to follow-up during the treatment period. There was no difference in the rate of recovery between the moclobemide-treated and the placebo-treated groups. Hence, all patients were evaluated as a single group for the purpose of this report. A neuropsychological test battery was performed, primarily to exclude patients with dementia, consisting of the following tests: Token test, Boston naming test, Rivermead Behavioural Memory Test (RBMT)/pictures, faces and orientation, memory, and 4 sub-tests of the Wechsler Adult Intelligence Scale-Revised-Neuropsychological Instrument, and apraxia tasks (8). Reasons for patients not performing the test were: aphasia too severe, patients too tired, or administrative problems. Patients with a low score (n = 3) on the date and orientation of the RBMT were classified as demented.

**Aphasia tests**

All aphasia tests were performed in random order by the same speech pathologist. Both the NGTA and the ANELT are well validated and are available in Swedish. Complete recovery was defined as full score on both the NGTA and the ANELT test, a Token test score of more than 30, and normal language in the opinion of both the patient and the speech pathologist.

**Standardized test.** The NGTA is based on the Boston terminology, is similar to the Western Aphasia Battery (WAB) (9), and measures fluency, comprehension, naming and repetition in addition to writing and reading. The sum of the total score of the main variables yields the aphasia coefficient (Coeff). Coeff is a measure of the severity of language impairment and constitutes the degree of aphasia. In study I, in which the aphasic patients were tested earlier after stroke onset than in study II, a shorter, adjusted version, representative of the entire NGTA was used (5). The short version takes 10–15 min to perform. In study II the entire NGTA was used, which takes 30–45 min. The Coeff has a range of 0–59 in study I and 0–217 in study II. The NGTA recognizes that some aphasic patients have a mixture of 2 or more aphasia syndromes. These patients are therefore classified as mixed non-fluent or mixed fluent. Fluent aphasias are: Wernicke, conduction, transcortical sensory, anomic and mixed fluent. Non-fluent aphasias are: global, Broca, transcortical motor and mixed non-fluent (4).

**Functional test.** ANELT is a measure of verbal communication ability and was used to assess the degree of aphasia. In this functional test the understandability of the patients’ message and the intelligibility of the utterance are each rated on a 5-point scale, where 1.0 indicates the most severe degree of aphasia. Each parallel test consists of 10 items, and takes a total of approximately 15 minutes to perform. The items are constructed as scenarios of familiar daily life situations,
Clinical and prognostic properties of aphasia tests

The test starts with 2 training tasks to ensure that the patient understands the idea of the test. The ANELT understandability score indicates the severity of the communication disability, and is a measure of the degree of aphasia. Since patients in study I with an initial score between 4.0 and 5.0 had a high rate of complete spontaneous recovery, such patients were excluded from study II. In study I, a score of 0 was given when the patient, due to severe aphasia, was incapable of taking instructions and/or giving an answer.

Other measures of communication and language ability

The Token test was used in study I to discriminate aphasia during follow-up (10). The version used has a range of 0–36, and the cut-off was set at 30 to define aphasia.

Neurological impairment was assessed according to the Scandinavian Stroke Supervision Scale (SSSS) (11). The study investigator (physician) performed this test at every visit. One item of the SSSS measures the degree of aphasia on a scale of 1–3, where 1 is normal. In study II this measure of the degree of aphasia was compared with the degree of aphasia measured by Coeff and ANELT.

Statistical methods

Data are presented as mean and standard deviation (SD), or as median with quartiles, as appropriate. A 5% probability ($p$) was considered statistically significant. Contingency tables were evaluated using the $\chi^2$ test. Statistical comparisons between groups were made using one-way analysis of variance. For the correlation between Coeff and ANELT, the best-fit regression line was expressed as an exponential second-degree polynomial equation. For multivariate regression analyses, multivariate analysis of variance was used. Analyses were carried out with the JMP, version 5.1 (SAS Institute Inc., Cary, NC, USA). Receiver operator characteristic (ROC) curves is a plot of the sensitivity to predict a complete recovery (1–sensitivity) for each value of the Coeff and ANELT, to give a graphical representation of the relationship between false-positive and true-positive rates. The relationship was evaluated as the area under the curve.

RESULTS

Correlations between the 2 aphasia tests

The degree of aphasia measured by Coeff and ANELT in studies I and II is shown in Table II, where the percentages of the maximum value are also given. Median values for Coeff and ANELT for all patients tested at both baseline and 6 months ($n = 72$ in study I, and $n = 76$ in study II) are shown in Fig. 2.

<table>
<thead>
<tr>
<th>Study</th>
<th>Baseline</th>
<th>3 months</th>
<th>6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td>ANELT</td>
<td>Coeff</td>
<td>ANELT</td>
</tr>
<tr>
<td>Study I</td>
<td>119</td>
<td>34</td>
<td>57</td>
</tr>
<tr>
<td>Study II</td>
<td>89</td>
<td>54</td>
<td>92</td>
</tr>
<tr>
<td>Study I</td>
<td>72</td>
<td>53</td>
<td>90</td>
</tr>
<tr>
<td>Study II</td>
<td>89</td>
<td>84</td>
<td>40</td>
</tr>
<tr>
<td>Study I</td>
<td>76</td>
<td>183</td>
<td>87</td>
</tr>
</tbody>
</table>

Coeff: Coefficient in Norsk Grunntest for Afasi (NGTA); ANELT: Amsterdam-Nijmegen Everyday Language Test. In study I a short version of NGTA was used, Coeff range 0–59. In study II the full version of NGTA was used, Coeff range 0–217. The ANELT test was the same in both studies. In study II patients with ANELT 4.0–5.0 at baseline were not included.

Fig. 2. Degree of aphasia measured with Coefficient (Coeff) in Norsk Grunntest for Afasi (NGTA) and Amsterdam-Nijmegen Everyday Language Test (ANELT) at baseline and at 6 months. Median values and interquartiles. The degree of aphasia improved for both Coeff and ANELT from baseline to 6 months in both study I and study II; $p < 0.0001$ for all.

Fig. 3. Relations between Coefficient (Coeff) in Norsk Grunntest for Afasi and Amsterdam-Nijmegen Everyday Language Test (ANELT) at baseline (left hand panel) and at 6 months (right hand panel) for study I and study II. The best-fit regression line was expressed as an exponential second-degree polynomial equation.
The correlation between the degree of aphasia according to Coeff and ANELT are presented in Fig. 3. For the individual patient, Coeff was sensitive to differentiate between patients with low values on ANELT, whereas ANELT was sensitive to differentiate between patients with high Coeff values.

**Prognostic capacity**

In study I, 24 aphasic patients (20%) recovered completely by 6 months. Among patients with Coeff ≥ 49 at baseline, 20 recovered completely, and among patients with ANELT ≥ 3.5 at baseline, 19 recovered completely. The latter 19 patients were all among the 20 patients who recovered completely.

Accordingly, Coeff and ANELT were equivalent in predicting complete recovery, as shown by the area under the curves in Fig. 4. Coeff with a sensitivity of 83% (corresponding to Coeff 49.0) had a specificity of 77% to predict complete recovery, while ANELT with a sensitivity of 79% (corresponding to ANELT 3.5) showed a specificity of 83%.

As shown in Fig. 2, the recovery rate was almost parallel between the 2 tests during the first 6 months in both studies. Recovery, measured as the difference between Coeff at baseline and at 6 months (Coeff 0–6), and the difference between ANELT at baseline and at 6 months (ANELT 0–6), was inversely related to the score at baseline in univariate analyses (p < 0.001 for both). Thus, the lower the initial score in either test, the higher the rate of recovery. In a multivariate analysis in study I including age, Coeff and ANELT, and fluency at baseline, Coeff and fluency were significant for Coeff 0–6 (p < 0.05), and ANELT and fluency significant for ANELT 0–6 (p < 0.001). Similar multivariate analyses in study II showed that age, ANELT, and Coeff correlated significantly with recovery measured by Coeff 0–6 (p < 0.05 for all), and age, ANELT, and Coeff correlated significantly with ANELT 0–6 (p < 0.05 for all). Increasing age correlated negatively with recovery (p < 0.05).

**Correlation between aphasia measures and other tests**

Neuropsychological tests could be carried out in 37 patients in study II. Multivariate analyses, including Coeff and ANELT at baseline, age, RBMT/picture and face, Block span total, forward, and backward, showed that Coeff at baseline, Block span total, forward, and backward, correlated with recovery, measured by Coeff 0–6 (p < 0.0001, p = 0.06, p < 0.01, and p = 0.07, respectively). With the same variables as above Coeff and ANELT at baseline, and RBMT/picture and face was correlated to recovery measured by ANELT 0–6 (p = 0.01, p < 0.0001, p = 0.01, and p = 0.05, respectively).

In study II the aphasia item in SSSS showed a modest correlation with Coeff and ANELT at baseline (r² = 0.39 and 0.33, respectively, p < 0.0001 for both). At 6 months, r² values were 0.52 and 0.53, respectively (p < 0.0001 for both).

**DISCUSSION**

The high and consistent correlation between NGTA and ANELT is interesting. In accordance with the International Classification of Impairments, Disabilities, and Handicaps (12), NGTA measures language impairment, while ANELT measures functional communication, which is considered a measure of disability even though ANELT is a verbal test. According to the more recent International Classification of Functioning, Disability and Health, both NGTA and ANELT are tests that involve body functions and activities (13). There is a difference between assessment of language impairment and how the individual gets along in a communicating world. Standardized aphasia tests focus on language deficit, while functional communication tests measure what the patient can do. ANELT has been used to assess functional communication, but our results show that ANELT in several respects is similar to Coeff in the NGTA test.

Another interesting finding is that Coeff in the standardized test could discriminate better among those with severe aphasia, i.e. ANELT 0 or 1, while the functional test, ANELT could discriminate better among those with a higher score according to the Coeff, as shown in Fig. 3. This indicates that NGTA is more suitable in the acute phase and ANELT in the subacute/chronic phase.

The Communicative Effectiveness Index (CETI) is a functional communication test assessed by significant others, which measures both verbal and non-verbal communication (14).

![Fig. 4. Receiver operating characteristic curves for Coefficient (Coeff) in Norsk Grunntest for Afasia and Amsterdam-Nijmegen Everyday Language Test (ANELT) show the sensitivity and specificity to predict complete recovery from aphasia. A Coeff value ≥ 49 gives 83% sensitivity and 77% specificity. An ANELT value ≥ 3.5 yields 79% sensitivity and 83% specificity. The area under the curve (AUC) summarizes the information contained in the curve and a higher value indicates a stronger correlation for Coeff and ANELT, respectively, to predict outcome.](image-url)
Aftonomos and colleagues showed that WAB and CETI have a strong positive correlation before treatment in chronic aphasic patients (15). Both tests could demonstrate improvement following treatment, but after treatment the correlation between the 2 tests reversed (15). Another study revealed strong and consistent correlations between WAB and CETI (16), similar to our findings with Coeff and ANELT. This study (16) was a trial with acute stroke patients with aphasia, i.e. similar to the subjects in the present study. The construction of the NGTA is similar to other standardized tests, while ANELT might be less functional than CETI, which is both verbal and non-verbal. Thus, it seems that improvements on the impairment and functional levels are more closely related in the acute and subacute stage, while in the chronic situation improvement on a functional level is more pronounced.

Many standardized and functional tests take approximately one hour to perform, which is too long for acute aphasic patients to endure. In our study, use of the short version of the NGTA was always feasible in the acute phase, as long as the patient was able to participate. The ANELT test, however, could not be performed in 1 out of 4 aphasic patients within the first week after stroke onset. ANELT is a more complex test and is easily misunderstood by patients in the acute phase. It requires more executive and abstraction ability. In study I the patients were tested sooner after the onset of stroke, and some were not alert enough to carry out the ANELT, and consequently scored 0 on this test.

Our findings show that patients who will recover completely can be identified with high accuracy, as displayed by the receiver operating characteristic curves in Fig. 4. The NGTA test had somewhat higher sensitivity (83% vs 79%) to predict complete recovery, while ANELT had higher specificity (83% vs 77%). Thus, a short version of a standardized aphasia test measuring the degree of aphasia seems to be sufficient for predicting outcome and prognosis. The predicting capacities for complete recovery up to 18 months were consistent with the findings at 6 months.

The item for assessment of aphasia in the neurological impairment scales is simple and quick to perform, and has been used to measure the degree of aphasia in many studies (17). However, in our study the correlations between this item and the degree of aphasia according to standardized and functional tests were only modest. In the acute stage there will be many false-positives cases, whereas patients with milder form of aphasia may be diagnosed as normal (false-negative) during follow-up. This was true in study II, and has also been shown by Thommessen et al. (18). The use of a neurological impairment scale to indicate the degree of aphasia is, accordingly, too imprecise, and is not sufficient as an outcome measure.

Our findings confirm previous observations that the possibility of improvement is greater with a more severe degree of initial aphasia (2, 3, 19, 20). Indeed, the improvement (as measured by Coeff or ANELT) was related to the initial degree of aphasia. However, these findings are due to the ceiling effect of the tests. Unfortunately, all currently available tests have ceiling effects. The rate of complete recovery depends on the time from the onset of stroke to the initial evaluation of aphasia. Burst et al. (3) reported complete recovery in 25% of aphasic survivors within the first 3 months. In study I, in which patients were tested within the first week after stroke onset, a good third of the survivors recovered completely by 18 months. In study II, in which patients with ANELT > 4.0 were excluded and the baseline test was performed within the first 3 weeks, none recovered completely. In the present investigation, complete recovery was defined as a full score on both NGTA and ANELT, a score on the Token test of more than 30, and normal language according to both the patient and the speech pathologist. An important determinant for prognosis is, consequently, the initial severity of the language impairment.

Neuropsychological tests often depend on verbal processing, and are therefore less adequate for patients with severe aphasia. Multivariate analysis, including neuropsychological tests, showed that working memory and visual memory were correlated to Coeff and ANELT scores, respectively. Cognitive assessment should be limited strictly and carefully to the functions most necessary for evaluating in the acute stage. A more extensive evaluation of cognitive function, if needed, should be performed in the subacute stage.

In conclusion, there is a close correlation between NGTA and ANELT in the assessment of the degree of aphasia. Both tests can predict complete recovery of aphasia in acute stroke patients with high sensitivity and specificity. The aphasia item in the neurological scales is insufficient for evaluation of the outcome of aphasia. The short version of NGTA can be performed easily in the acute stage. A standardized test appears to be more suitable for patients with aphasia in the acute stage, while a functional test is more suitable in the subacute/chronic stage.

ACKNOWLEDGEMENTS

We are grateful to Dr Peter Borenstein, Skene, Sweden for the Swedish versions of the Norsk Grunntest for Afasi and the Amsterdam-Nijmegen Everyday Language Test, and for fruitful suggestions and advice.

This study was supported by the Swedish Stroke Association Foundation and Funds, the Seratiner Hospital Foundation, and Karolinska Institutet.

REFERENCES