Research Project Report

(PSY9188)

University of Lincoln

School of Health, Life and Social Sciences

Doctorate in Clinical Psychology

2009

The Validity of the Addenbrooke's Cognitive Examination- Revised (ACE-R) in Acute Stroke

Katie J. Morris

Submitted in part fulfillment of the requirements for the Doctorate in Clinical Psychology

Title

The Validity of the Addenbrooke's Cognitive Examination- Revised (ACE-R) in Acute Stroke

Contents

Section 1

Abstract

Background: The MMSE is commonly used as a screening instrument for cognitive impairment in stroke services. However, recent research has shown that it has poor diagnostic validity for use in this patient population. The purpose of this study was to examine the validity of the ACE-R as an alternative screening measure for use in stroke.

Objectives: The first objective was to determine whether the ACE-R is more accurate than the MMSE at detecting overall cognitive impairment in stroke. The second objective was to determine the accuracy of the ACE-R subscales for detecting impairments in specific cognitive domains.

Methods: This study had a cross-sectional design. 40 patients were recruited from an inpatient stroke service. They were administered the ACE-R (which includes the MMSE), and a battery of more detailed neuropsychological tests, which served as the 'gold standard' for classification of impairment. The diagnostic validity of the ACE-R and MMSE was determined by ROC analysis.

Results: Both the MMSE and the ACE-R were found to have inadequate diagnostic validity for the detection of overall cognitive impairment. No cut-scores scores could be identified which yielded test sensitivity of >80% and specificity of >60%. Levels of specificity were particularly poor. The ACE-R subscales showed a similar pattern of performance, indicating inadequate validity for the detection of impairment in specific areas of cognitive functioning.

Conclusions: There was no support for the use of the MMSE or the ACE-R when screening for cognitive impairment in acute stroke. Further research should focus on the identification of an alternative measure.

Key words: Stroke, cognitive impairment, screening, sensitivity, specificity.

Section 2

Statement of contribution

The ACE-R stroke validation project was initially the concept of two Clinical Psychologists affiliated with Nottingham University Hospitals NHS Trust Clinical Psychology service; Professor Nadina Lincoln and Dr Vicki Hacker. The Trainee (author) contributed to the project design and independently carried out the literature review and application for ethical approval. The recruitment of patients was conducted by the Trainee, assisted by an Assistant Psychologist (Daniel Stark) and an MSc student (Kimberly Fletcher). The collection of patient demographic information, and the administration and scoring of the ACE-R screening tests was carried out by the Assistant Psychologist and MSc student. The Trainee was responsible for the administration, scoring and interpretation of the full neuropsychological test batteries. Clinical reports on each patient were written for the Stroke Service by the Trainee, under the supervision of Dr Vicki Hacker, Clinical Psychologist. Data entry was completed by the Trainee, assisted by the MSc student. Data analysis was independently carried out by the Trainee.

Section 3

Research paper

Introduction

Cognitive impairment is common following stroke, and is present in approximately 70% of patients in the acute stages of recovery [1]. The nature and severity of cognitive impairment varies according to the location of the stroke, but may include problems with memory, perception, language, attention, or executive functioning [1,2,3]. Over and above the physical impairments caused by stroke, cognitive difficulties impact on a person's ability to carry out tasks of everyday living and affect their ability to live independently [4,5]. As such, there is a clear need to accurately detect the presence of cognitive impairment in acute stroke, in order to inform prognosis, rehabilitation planning, and provision of care. Accordingly, the National Clinical Guidelines for Stroke, recommends that all stroke patients should be assessed for the presence of cognitive impairment [6]. (See appendix 1.1 for further information regarding cognitive impairment in stroke.)

Detailed neuropsychological testing is time consuming and costly for services, and tiring for acutely unwell patients. Thus, brief cognitive screening measures are required to identify those patients with impairment and to highlight problem areas for further assessment. A screening measure should have high levels of *sensitivity* (able to correctly identify people with problems) and high levels of *specificity* (able to correctly identify people without problems). Where screening pre-empts full testing, it may be better to compromise specificity for sensitivity to ensure that most of those people with problems are correctly identified. It is recommended in the literature that screening measures have a sensitivity of 80-

90% and a specificity of greater than 60% [7]. In addition to their psychometric properties, screening measures should also be brief, easy to administer by staff not trained in neuropsychological assessment, and suitable for bedside (hospital) administration. (See appendix 1.2 for further information on the principles of cognitive screening.)

The Mini Mental State Examination (MMSE) [8] is the most widely used of all cognitive screening measures, in stroke and other clinical services [9]. Like most screening measures it was originally developed to screen for dementia in community geriatric populations, consisting mainly of patients with Alzheimer's type dementia. It has been found to have high levels of diagnostic validity for this purpose, and is especially accurate at detecting moderate-severe impairment [10]. However, recent research has indicated that, in acute stroke populations, the MMSE has poor diagnostic validity for the detection of cognitive impairment [1,11,]. The most recent paper reported that, in stroke, the MMSE has a sensitivity of 34%, a specificity of 70%, and Area Under the Curve (AUC) = 0.67, p=0.13) [1]. The diagnostic validity is likely to be so poor in stroke due to this population comprising of patients with both focal deficits as well as dementia syndromes. The MMSE only measures a limited range of cognition, focussing on verbal functions and omitting any measure of executive functioning. As such, patients suffering predominantly from impairments in executive or perceptual functioning are missed by the MMSE [1]. (See appendix 1.3.1 for further information on the MMSE.)

Other commonly used screening tests are the Cambridge Cognitive Examination (CAMCOG) [12], and a shortened version, the Rotterdam-Cambridge Cognitive Examination (R-CAMCOG) [13]. These were also designed for use in community dementia populations. They offer more comprehensive assessment compared to the MMSE and both have been shown to have superior diagnostic validity for the detection of dementia syndromes, including vascular dementia [13,14,15]. However, their utility for the detection of cognitive impairment in an acute stroke population (patients with focal impairments and patients with vascular dementia) has not been examined. Their potential for this purpose may be limited due to the omission of a measure of executive functioning. Furthermore, the CAMCOG/ CAMCOG-R are likely to be impractical for use in acute stroke as they require specialised test equipment. (See appendix 1.3.2 and appendix 1.3.3 for further information on the CAMCOG and R-CAMCOG.)

The Addenbrooke's Cognitive Examination-Revised (ACE-R) [16], successor to the Addenbrooke's Cognitive Examination (ACE) [17], is an alternative option. Like the MMSE and the CAMCOG, the ACE-R was also designed for use in community dementia populations, but its design and psychometric properties suggest it may be more suitable for use in acute stroke. The ACE-R covers a more comprehensive range of cognitive impairment than both the MMSE and CAMCOG, including a measure of executive functioning. A further advantage of the ACE-R is that it provides normative data for five subscales, allowing accurate analysis of the pattern of cognitive impairment. Unlike the CAMCOG tests, the ACE-R does not require any materials other than the test sheet to

administer. Validity studies have indicated that the ACE-R has increased sensitivity and specificity for the detection of dementia, compared to the MMSE, although there has been some debate over the ideal cut-off score [16, 18]. One study has also given early indications that the ACE-R is a valid measure for the detection of cognitive impairment in brain injury [19]. (See appendix 1.3.4 and appendix 1.3.5 for further information on the ACE and ACE-R.)

The diagnostic validity of the ACE/ACE-R for the detection of cognitive impairment in acute stroke patients has not yet been examined, and this was the aim of the current study. There were two main research objectives. The first objective was to determine whether the ACE-R is more accurate than the MMSE at detecting the presence of overall cognitive impairment. The second objective was to determine the accuracy of the ACE-R in detecting impairments in specific areas of functioning (memory, attention, perception and executive functioning). In relation to the first objective, it is hypothesised that the ACE-R will be more accurate at detecting the presence of cognitive impairment compared to the MMSE in stroke. This prediction is made on the basis of the results found in dementia populations [17], and the assumption that the ACE-R is a more comprehensive assessment. It is not possible to make a prediction of how good the ACE-R is at identifying impairment in specific domains as this has not previously been investigated. However, positive results might be expected given the moderate-to-good correlations between scores on the ACE subscales and related neuropsychological tests [17].

Methods

Participants were recruited from the stroke service at Nottingham City Hospital. This service comprises of one hyper-acute stroke ward and 3 acute stroke wards. Patients were identified by examination of the medical notes held on the wards. The main inclusion criterion into the study was a confirmed diagnosis of stroke, determined by a CT scan. Patients were excluded from the study if they had a history of psychiatric problems or if they were blind, deaf, or too ill or drowsy (determined in consultation with the ward nurses). Patients were also excluded if they did not speak English, or if they had moderate or severe aphasia as tested by the Sheffield Screening Test for Acquired Language Disorders [20] (see appendix 1.4), scoring below the age-related cut-off scores stated in the manual. These patients would have been unable to complete the cognitive assessments. Demographic information was obtained from the medical records for each eligible patient, comprising; age, gender, date of stroke, type and hemisphere of stroke (determined by CT scan), presence of hemiparesis (determined by medical examination), presence of hemianopia (determined by the Ophthalmologist), and functional ability (determined by the Barthel Index [21], administered by an Occupational Therapist).

Patients who were eligible for the study were first administered an ACE-R screening test (which includes the MMSE). This was carried out mainly by a Research Assistant or Assistant Psychologist, although in some cases, patients had already been administered an ACE-R by a junior doctor, in which case a copy was retrieved from the medical file. Up to this point, informed consent for

participation in the study was not sought as the ACE-R was carried out as part of routine clinical practice. After patients had received an ACE-R they were approached for full neuropsychological testing and, at this time, informed consent for the study was sought. Each participant received a battery of 'gold standard' neuropsychological tests, selected for their proven validity and appropriateness for use with neurological patients (see appendix 1.5). The neuropsychological test battery was administered in a consistent order comprising of; the *Logical Memory* subtest from the Wechsler Memory Scales III (WMS III) [22] to assess verbal memory, the *Rey-Osterrieth Complex Figure test* (recall) [23] to assess visual memory, the *Star Cancellation test* from the Behavioural Inattention Test [24] to assess visual neglect, the *Rey-Osterrieth Copy* task [23] to assess visuo-spatial perception, the *Hayling Sentence Completion test* [25] and a *Verbal Fluency* test (F,A,S) [26] to test executive functioning, and the *Letter-Number Sequencing* and *Digit Span* subtests from the WMS-III as measures of attention. Neuropsychological testing lasted approximately one hour per participant, and was carried out by a Trainee Clinical Psychologist who was blind to the participants' ACE-R scores. The tests were scored and interpreted by the Trainee Clinical Psychologist, under the supervision of a qualified Clinical Psychologist. The majority of tests were scored according to the scoring procedures and norms presented in each test manual. The exceptions were the Hayling Sentence Completion and F,A,S tests which were interpreted using alternative norms, derived from populations with a greater number of older adults [27,28]. Percentile scores for each test were derived using the published means and standard deviations. Patients were classified as impaired on a test if they scored below the fifth percentile.

The exception to this was if it was deemed likely that a patient failed a particular test due to a different cognitive problem or a non-cognitive problem (in this case the test was classed as void). Following interpretation, the results were fed back to patients verbally, and patients were given information regarding the nature of their specific impairments and compensation strategies. Written reports were provided for the medical team and filed in the notes. (The recruitment procedure is summarised in the diagram in Appendix 1.6.)

Data was analysed using SPSS version 14. Demographic data and descriptive data for the test results were calculated. Next, analysis focussed on determining the diagnostic validity of the MMSE and ACE-R for detecting overall cognitive impairment at published cut-offs of 27 and 24 (MMSE), and 88, 82 and 75 (ACE-R). Sensitivity and specificity was determined by comparison with the presence of cognitive impairment as indicated by the neuropsychological test battery (impairment in any one or more of the tests). Positive and negative predictive values were calculated to determine the accuracy of the tests taking into account base rates of cognitive impairment in stroke. Receiver Operating Curves were also plotted to summarise the overall hit rate of diagnoses (i.e. the probability that the result for a randomly chosen positive case will exceed the result for a randomly chosen negative case). Following this, analysis of validity focused on the ACE-R individual subscales. This was determined by similar methods, comparing impairment on ACE-R subscales with impairment on tests measuring equivalent cognitive domains in the neuropsychological test battery. Performance on the ACE-R Attention and Orientation subscale was compared to the presence of impairment on either on WAIS-III Digit Span and/ or WAIS-III

Letter-Number Sequencing. The Memory subscale was compared to presence of impairment on either WAIS-III Logical Memory and/or Rey Figure Recall). The Fluency subscale was compared to presence of impairment on either FAS and/or impairment on the Hayling Sentence Completion test (indicated by failure on part two of the test, in the absence of failure on part one which controls for word finding difficulties and processing speed). Finally, the Visuospatial subscale was compared to impairment on either Star Cancellation and/or Rey Figure Copy. The ACE-R language subscale was not examined as the presence of aphasia had been used as one of the exclusion criteria.

Ethical approval for the study was obtained by the National Research Ethics Service, Nottingham Research Ethics Committee 1. No significant risks or adverse effects were identified for research participants in the study. Minor or routine ethical issues, such as informed consent, were managed appropriately, and approved by the Research Committee (see Appendix 1.7.)

Results

There were 64 stroke patients who were identified as meeting the inclusion and exclusion criteria for the study. All these patients completed the ACE-R/MMSE screen as part of routine clinical care. However 24 (34%) patients did not go on to complete the full neuropsychological test battery; 22 (92%) of these patients were discharged from hospital before they could be tested, it was discovered one patient was illiterate and another became unwell during testing. Thus the final number of patients recruited into study was 40. T-tests and chi-square analysis showed that there were no significant differences between the recruited and non-recruited patients in terms of demographic variables or performance on the screening tests (see appendix 2.1 and 2.2).

Demographic information

Of the 40 patients recruited to the study, 21 were male and 19 were female. The median age was 78 years (IQR 15.5), and the median years of education was nine years (IQR 1). Fifteen (38%) patients had Lacunar stroke (LACS), seven (17.5%) had Total Anterior Circulation Stroke (TACS), eight (20%) had Partial Anterior Circulation Stroke (PACS), four (10%) had Posterior Circulation Stroke (POCS) and type of stroke was unknown for six (15%) patients. Twenty-four (60%) patients had right hemisphere stroke, 14 (35%) had left hemisphere stroke, and hemisphere of stroke was unknown for two (5%) patients. Hemiparesis was present in 18 (45%) patients, absent in 20 (50%) patients and unknown for two (5%) patients. Hemianopia was present in 13 (33%) patients, absent in 25 (63%) patients and unknown for 2 (5%) patients. The median

score on the Barthel Index at recruitment was nine (IQR 5.5). The median interval between stroke and receiving the screening measure was 14 days (IQR 25.3). The median interval between receiving the screening measure and the neuropsychological test battery was two days (IQR 6). The median interval between stroke and receiving the neuropsychological test battery was 18 days (IQR 27.5).

Descriptive data

Table 1 shows the descriptive results from the screening tests and the neuropsychological test battery. The number of patients impaired on the MMSE and ACE-R varied according to the cut-off score used. According to the neuropsychological test battery, 31 (78%) patients were found to have cognitive impairment in at least one cognitive domain. The most common type of impairment found was perceptual impairment, with 23 (61%) patients impaired on either the Star Cancellation or Rey Figure Copy task. Executive impairment was almost as common, with 22 (60%) patients impaired on either the FAS or Hayling Sentence Completion task. A total of 14 (47%) patients had impairment on either the verbal or visual memory recall tasks, and only four (11%) patients had impairment on the tasks of attention (Digit Span or Letter-Number Sequencing).

Analysis showed that performance on the MMSE, ACE-R and the neuropsychological test battery was not significantly related to age, gender, hemisphere of stroke, presence of hemiparesis, presence of hemianopia, years of education, score on the Barthel Index, and the number of days between

stroke and administration of test. However, there was a significant relationship between type of stroke (LACS, TACS, PACS, and POCS) and performance on the MMSE, $F(3,30)= 5.5$, $p< 0.05$, and performance on the ACE-R, $F(3,30)= 4.16$, $p<0.05$. The relationship between type of stroke and impairment on the neuropsychological test battery approached significance levels, $x^2(3, N=34)= 6.94$, $p=0.07$. (See appendix 2.1 and 2.3 for full description of analysis).

Validity analysis

First, the diagnostic validity of the MMSE and ACE-R was examined in relation to the presence of overall cognitive impairment as classified by the neuropsychological test battery (impairment in one area of cognitive functioning). (See Appendix 2.4 for crosstabulations of impairment on the screening tests and the neuropsychological test battery.) Sensitivity and specificity values for the ACE-R and MMSE (at the published cut-offs) are reported in Table 2. Alongside, are corresponding positive and negative predictive values for the base rate of cognitive impairment reported in acute stroke [1]. Table 2 shows that none of the published cut-off scores give both adequate levels of sensitivity (>80%) and specificity (>60%). A cut-off of 27 on the MMSE and 82 on the ACE-R are, arguably, the least poor. On the MMSE, a cut-off of 27, gives good rates of sensitivity (81%), although unsatisfactory levels of specificity (22%). On the ACE-R, a cut-off of 82, gives fair levels of sensitivity (77%), but unsatisfactory specificity (44%). The positive and negative predictive values were very similar for both tests at all cut-offs. Positive predictive values were between 70-80%, indicating reasonable accuracy for positive results, but negative predictive values were all below 50%,

indicating poor accuracy for negative results. Further analysis revealed that no alternative cut-offs provided suitable levels of specificity (>80%) and sensitivity (>60%). (See Appendix 2.5 for sensitivity and specificity rates at a full range of cut-offs.) Receiver operating curves are presented in Figure 1. The area under the curve (AUC) indicates that, overall, both tests have inadequate diagnostic accuracy, performing no better than chance (MMSE AUC= 0.57, $p> 0.05$; ACE-R AUC= 0.58, $p> 0.05$).

Secondly, the diagnostic validity of four ACE-R subscales was examined in relation to the presence of impairment on tests measuring equivalent cognitive domains in the neuropsychological test battery. (See Appendix 2.6 for crosstabulations of impairment on the ACE-R subscales and cognitive domains measured by the neuropsychological test battery.) Sensitivity and specificity for each ACE-R subscale are presented in Table 3. These were calculated according to the cut-offs published in the ACE-R manual, which varies according to age of each patient. For all subscales, sensitivity is good (>80%) and for the Attention and Orientation subscale, sensitivity is excellent (100%). Specificity is unsatisfactory for all subscales (<60%). Further analysis revealed that no alternative cut-offs could be found for any subscale which gave a specificity of >60% whilst retaining sensitivity of >80%. (See Appendix 2.7 for sensitivity and specificity rates for each subscale at a full range of cut-offs.) Receiver operating curves are presented in Figure 2. Although all subscales had consistently poor levels of specificity, the Area Under the Curve statistic indicated that, overall, using most of the subscales to predict impairment in specific areas was better than guessing; Attention and Orientation (AUC=0.84,

$p < 0.05$), Fluency (AUC=0.72, $p < 0.05$) and Visuospatial (0.71, $p < 0.05$). The memory subscale was the exception, performing statistically no better than chance (AUC=0.58, $p > 0.05$).

Discussion

In comparison with the detailed neuropsychological battery of tests, neither the ACE-R nor the MMSE could detect the presence of cognitive impairment with adequate levels of sensitivity, as well as specificity. At the higher published cut-off scores, both tests yielded fair-good levels of sensitivity, but levels of specificity were inadequate, yielding high numbers of false positive diagnoses. No alternative cut-offs could be identified which increased specificity without compromising sensitivity. Area Under the Curve statistics indicated that, overall, the diagnostic accuracy of the tests was not much better than chance, meaning the diagnosis would be wrong in nearly 50% of cases. Given the high base rates of cognitive impairment in acute stroke (70%), it is therefore more accurate to assume all patients have cognitive impairment, without even doing any tests, because you would only misdiagnose 30% of cases.

Compared to performance on individual cognitive domains of the neuropsychological test battery, none of the ACE-R subscales examined could identify impairment with both adequate sensitivity and adequate specificity. All subscales gave good-excellent sensitivity at the published cut-offs, but specificity was unsatisfactory, giving high rates of false positives. No suitable alternative cut-offs were identified. Area Under the Curve statistics indicated that, overall, the diagnostic accuracy of the scales was adequate, with the exception of the memory scale. We can thus retain only limited confidence that the ACE-R subscales can accurately assist in identifying the pattern of cognitive impairment in stroke.

The failure of both tests in this study appears to lie with poor specificity and, therefore, high rates of false positive diagnoses. These findings are contrary to what was predicted. Given previous literature, it was correctly predicted that the MMSE would perform poorly in a stroke population [1]. However, it was expected the ACE-R would perform better than the MMSE given its superior performance in related fields [16,18], and its wider coverage of cognitive domains. It is possible that certain characteristics of an acute stroke population have led to both screening tests having poor specificity compared to in a dementia population. Factors such as physical ill health and the anxiety of being in hospital, may have resulted in patients finding the MMSE/ACE-R too difficult and caused them to fail for reasons other than cognitive impairment. Such factors may have impacted particularly on the orientation, attention and memory sections which are heavily weighted in the MMSE/ACE-R. Unfortunately, this does not explain why previous studies of the MMSE in acute stroke have found poor sensitivity, rather than poor specificity [1,11]. Therefore, it is useful to consider how methodological shortcomings in the current study may have contributed to findings, especially the contrary findings of poor specificity.

One limitation of this study is the interval between the administration of the screening tests and the neuropsychological test battery (median= 2 days (IQR 6). In a population where recovery is rapid, this may have resulted in relatively better performance on the neuropsychological test battery, leading to incorrect false positive diagnoses on the screening tests (poorer specificity). However,

the fact that performance on screening and neuropsychological tests was not related to number of days since stroke, makes this hypothesis less likely to be true. Also, whilst one of the previous studies of the MMSE in stroke had no time lag between the MMSE and test battery [1], the other study [11], had a time lag of up to three months and still found the MMSE to have good specificity.

A second limitation of the current study is the relatively small sample size (40) and in particular, the small proportion of non-impaired patients (nine) compared to impaired patients (31). In previous studies investigating the MMSE in stroke [1,11], and the ACE/ACE-R in dementia [16,17], samples were much larger and included a better mix of impaired/non-impaired patients. In fact, most studies [1,16,17] included a non-impaired control group for comparison. Therefore, it is possible that the small sample size and imbalance of non-impaired/impaired patients in the current study may have lead to inadequate statistical power and falsely inflated rates of poor specificity.

A third limitation is that three of the nine patients who were classified as unimpaired by the neuropsychological test battery, did not manage to complete all tests due to non-cognitive problems such as poor hearing or onset of feeling unwell. We cannot be 100% certain these three patients would not be impaired on the tests they did not complete. These patients may have incorrectly contributed to high false positive classifications on the screening tests.

A forth limitation to the study is a possible recruitment bias due to the high attrition rates between the administration of the screening tests and the

neuropsychological test battery. This may have resulted in a final sample which is poorly representative of an acute stroke population, with a bias of more impaired patients. However, the proportion of impaired patients in this study (78%) is similar to that found previously (70%) [1], making this less likely. Furthermore, analysis indicated that, in terms of demographic variables or performance on the screening tests, there were no significant differences between patients who completed the battery and those who did not.

A final point worth considering is whether, rather than the MMSE/ACE-R lacking in specificity, the neuropsychological test battery was lacking in sensitivity. One area of failure in the battery may have been the tests chosen to measure attention, i.e., Digit Span and Letter-Number-Sequencing. These tests only measure one aspect of attention, working memory, and do not measure other facets of attention, i.e. focused attention, sustained attention, selective attention, and divided attention. Similarly, another area of weakness in the battery may have been in the tests chosen to measure executive functioning (Hayling Sentence Completion and F,A,S). Although these are well validated measures, they only really measure one aspect executive functioning which is cognitive flexibility/selective inhibition. There are other aspects of executive functioning, e.g. planning ability and abstract thinking, which the test battery did not measure. Further to this, the test battery did not include any measure of language. This was a deliberate decision as language impairment, measured on the Sheffield Screening Test, had been an exclusion criterion, in order that language did not confound performance on other tests. However, patients were only excluded if they had moderate-severe aphasia and it is possible that

patients with less severe language impairments were included in the study. The test battery would, therefore, have been relatively insensitive to these impairments, especially in comparison to the ACE-R which includes a large language component within the total score.

See Appendix 3.1 for discussion on how limitations could be addressed.

To summarise, the current study found both the MMSE and the ACE-R to have inadequate validity for the detection of cognitive impairment in stroke, particularly demonstrating poor levels of specificity. The ACE-R subscales showed a similar pattern of performance. Poor specificity of the screening measures in stroke may be due to certain characteristics of this population causing failure on the tests for reasons other than cognitive impairment. Given some limitations in the current study, including a small sample size, future research may aim to address limitations and replicate these findings. Further to this, future research should focus on the identification of an alternative screening measure, or adaptation of an existing measure for suitability in stroke. In the meantime, there remains no support for the use of the MMSE or the ACE-R when screening for cognitive impairment in acute stroke.

(See appendix 3.2 for further discussion on the implications of this research for Clinical Psychology services in stroke. See appendix 3.4 for reflections on the epistemological positioning of this research).

Key points:

- There was no support for the use of the MMSE or the ACE-R when screening for cognitive impairment in acute stroke.
- Further research is needed to identify an alternative screening measure with adequate levels of sensitivity and specificity.

References

1. Nys, G., van-Zandvoort, M., de-Kort, P., Jansen, B., Kappelle, L. & De-Haan, E. (2005). Restrictions of the Mini-Mental State Examination in acute stroke. *Archives of Clinical Neuropsychology, 20*(5), 623-629.

2. Nys, G.M.S., van-Zandwoort, M.J.E., de-Kort, P.L.M., Jansen, B.P.W., de-Haan, E.H.F., & Kapelle, L.J. (2007). Cognitive disorders in acute stroke: prevalence and clinical determinants. *Cerebrovascular Diseases, 23*(5-6), 408-416.

3. Engstad, T., Viitanen, M., & Almkvist, O. (2007). Cognitive impairment after stroke- diagnosis and management. *Tidsskrift for den Norske Legeforening, 127*(10), 1390-1393.

4. Claesson, L., Linden, T., Skoog, I., & Blomstrand, C. (2005). Cognitive impairment after stroke - impact on activities of daily living and costs of care for elderly people. *Cerebrovascular Diseases, 19*(2), 102-109.

5. Nys, G.M.S., van-Zandvoort, M.J.E., de-Kort, P.L.M., van-der-Worp, H.B., Jansen, B.P.W., Algra, A., de-Haan, E..H.F., & Kappelle, L.J. (2005). The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. *Neurology, 64*(5), 821-827.

6. Royal College of Physicians of London. (2004). *National Clinical Guidelines for Stroke* (2nd ed.) [brochure]. London: Authors.

7. Lincoln, N.B., Nicholl, C.R., & Flannaghan, T. (2003). The validity of questionnaire measures for assessing depression after stroke. *Clinical Rehabilitation, 17*, 840-846.

8. Folstein, M., Folstein, S., & McHugh, P. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 78-85.

9. Shulman, K.I., Hermann, N., Brodaty, H., et al. (2006). IPA survey of brief cognitive screening instruments. *International Psychogeriatrics, 18*, 281–94.

10. Tombaugh, T.N., & McIntyre, N.J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society, 40*(9), 922-935.

11. Blake, H., McKinney, M., Treece, K., Lee, E., & Lincoln, N. (2002). An evaluation of screening measure for cognitive impairment after stroke. *Age and Ageing, 31*(6), 451-456.

12. Huppert, F., Brayne, C., Gill, C., Paykel, E., & Beardsall, L. (1995). CAMCOG- a concise neuropsychological test to assist dementia diagnosis:

Socio-demographic determinants in an elderly population sample. *British Journal of Clinical Psychology*, 34, 529-541.

13. De-Koning, I., Dippel, D., van-Kooten, F., & Koudastall, P. (2000). A short screening instrument for post stroke dementia: the R-CAMCOG. *Stroke*, *31*(7), 1502-1508.

14. De-Koning, I., van-Kooten, F., Dippel, D., van-Harskamp, F., Grobbee, D., Kluft, C., & Koudstaal, P. (1998). The CAMCOG: a useful screening instrument for dementia in stroke patients. *Stroke, 29*(10), 2080-2086.

15. De-Koning, I., van-Kooten, F., Koudstaal, P.J., & Dippel, D.W.J. (2005). Diagnostic value of the Rotterdam-CAMCOG in post-stroke dementia. *Journal of Neurology, Neurosurgery and Psychiatry, 76*(2), 263-265.

16. Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21*(11), 1078-1085.

17. Mathuranath, P., Nestor, P., Berrios, G., Rakowicz, W., & Hodges, J. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology, 55*(11), 1613-1620.

18. Larner, A.J. (2007). Addenbrooke's Cognitive Examination-Revised(ACE-R) in day-to-day clinical practice. *Age Ageing, 36,* 685–6.

19. Gaber, T.A.Z.K. (2008). Evaluation of the Addenbrooke's Cognitive Examination's validity in a brain injury rehabilitation setting. *Brain Injury, 22*(7),589-593.

20. Syder D, Body R, Parker M & Boddy M (1993). *Sheffield Screening Test for Acquired Language Disorders*. Windsor: Nfer-Nelson.

21. Mahoney FI*, &* Barthel D. (1965). Functional evaluation: the *Barthel* Index. *Maryland State Medical Journal,14,* 56-61.

22. Wechsler, D. (1999). *Wechsler Memory Scale, third edition (WMS-III)*. Oxford: Harcourt Assessment.

23. Meyers, J., & Meyers, K. (1995). *Rey Complex Figure Test and Recognition Trial (RCFT)*. Oxford: Harcourt Assessment.

24. Wilson, B., Cockburn, J., & Halligan, P. (1987). Behavioural Inattention Test (BIT). Oxford: Harcourt Assessment.

25. Burgess, P.W. & Shallice, T. (1997). *Hayling and Brixton Tests*. Oxford: Harcourt Assessment.

26. Berg, E. (1948). *A* simple objective treatment for measuring flexibility in thinking. *Journal of General Psychology, 39*, 15-22.

27. Bielak, A.M., Mansueti, L., Strauss, E., & Dixon, R.A. (2006). Performance on the Hayling and Brixton tests in older adults: Norms and correlates. *Archives of Clinical Neuropsychology, 21*(2), 141-149.

28. Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*, 167-177.

Tables and figures

Table 1. Descriptive data from the screening tests and neuropsychological assessments

Assessments	*n*	Mean	SD	Range	Impaired n	Impaired %
Screening tests						
MMSE	40	24.4	3.6	17-30		
cut-off= 27					32	(80%)
cut-off= 24					19	(48%)
ACE-R	40	70.9	14.0	38-88		
cut-off= 88					40	(100%)
cut-off= 82					29	(73%)
cut-off= 75					23	(58%)
ACE-R Attention and Orientation	40	15.3	2.2	9-18	26	(65%)
ACE-R Visuospatial	39	11.7	4.0	0-16	29	(74%)
ACE-R Fluency	40	7.1	2.9	3-12	30	(75%)
ACE-R Memory	40	15.7	4.2	5-25	31	(78%)
Neuropsychological assessments (Overall)	-	-	-	-	31	(78%)
Star cancellation	38	49.3	10.3	7-54	9	(24%)
Rey Figure Copy	37	22.0	11.4	0.5-36	22	(60%)
FAS	40	27.2	11.6	6-55	4	(10%)
Hayling Sentence Completion (errors)	36	8.6	4.7	0-15	20	(56%)
Logical Memory- immediate recall	40	24.3	10.4	6-44	9	(23%)
Logical Memory- delayed recall	40	10.7	6.4	0-24	10	(25%)
Rey Figure- immediate recall	26	7.3	5.0	0-19	10	(39%)
Rey Figure- delayed recall	26	7.4	5.0	0-19	10	(39%)
Digit Span	40	15.8	3.4	6-23	0	(0%)
Letter-Number Sequencing	38	7.2	3.0	2-14	4	(11%)

MMSE= Mini-Mental State Examination
ACE-R= Addenbrooke's Cognitive Examination-Revised

Table 2. Validity of the MMSE and ACE-R for detecting overall cognitive impairment

Test/cut-off	Sensitivity	Specificity	PPV	NPV (base rate=70%)
MMSE				
27	81%	22%	71%	33%
24	52%	67%	78%	37%
ACE-R				
88	100%	0%	70%	-
82	77%	44%	77%	46%
75	58%	44%	71%	31%

Acceptable levels of sensitivity and specificity are highlighted

MMSE= Mini-Mental State Examination
ACE-R= Addenbrooke's Cognitive Examination-Revised

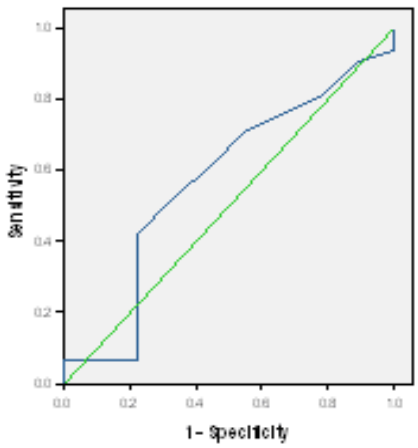
Table 3. Validity of the ACE-R subscales

ACE-R subscale/ cut-offs (by age group)	Sensitivity	Specificity
Attention and Orientation (17,17,16)*	100%	39%
Memory (18,19,17)*	85%	31%
Fluency (9,8,9)*	82%	36%
Visuospatial (15,14,14)*	87%	46%

Acceptable levels of sensitivity and specificity are highlighted

ACE-R= Addenbrooke's Cognitive Examination-Revised
* cut-offs applied to each patient according to age (50-59 years, 60-69 years,70+ years respectively).

ROC curve for MMSE and overall cognitive impairment



ROC curve for ACE-R and overall cognitive impairment

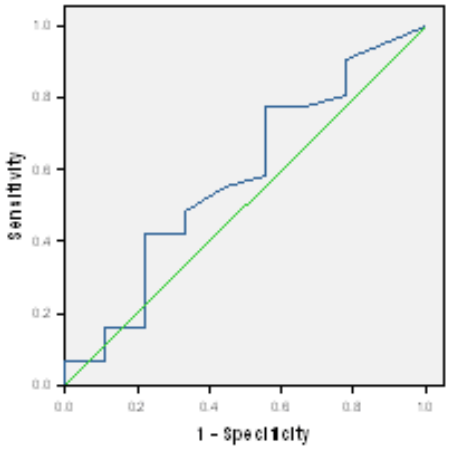
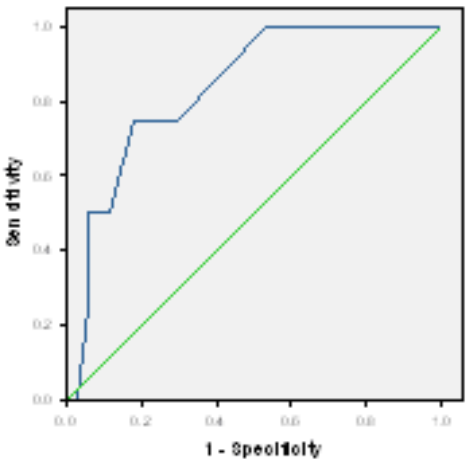
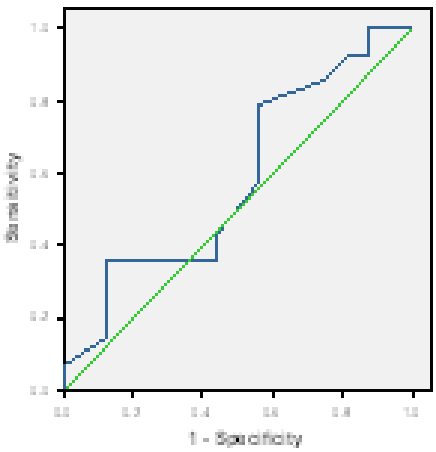


Figure 1. Comparisons between the screening tests and overall cognitive impairment on the neuropsychological test battery

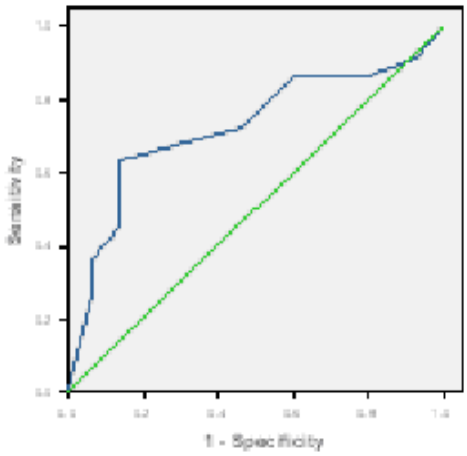
ROC curve for ACE-R 'Attention and Orientation' and attention



ROC curve for ACE-R 'Memory' and memory



ROC curve for ACE-R 'Fluency' and executive functioning



ROC curve for ACE-R 'Visuospatial' and perception

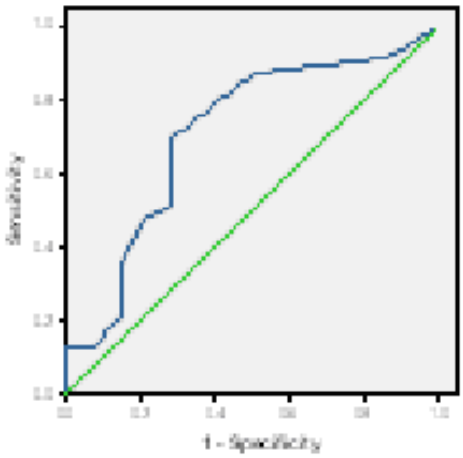


Figure 2. Comparison between ACE-R subscales and impairment on cognitive domains on the neuropsychological test battery

Section 4

Journal guidelines for authors

Guidelines for authors submitting to the journal 'Age and Ageing'

The research paper is written with a view to submit for publication to the Journal of Age and Ageing. This journal stipulates papers should be written according to the uniform requirements for manuscripts submitted to biomedical journals. These guidelines are presented below:

Abstract and Key Words

An abstract (requirements for length and structured format vary by journal) should follow the title page. The abstract should provide the context or background for the study and should state the study's purposes, basic procedures (selection of study subjects or laboratory animals, observational and analytical methods), main findings (giving specific effect sizes and their statistical significance, if possible), and principal conclusions. It should emphasize new and important aspects of the study or observations.

Because abstracts are the only substantive portion of the article indexed in many electronic databases, and the only portion many readers read, authors need to be careful that abstracts reflect the content of the article accurately. Unfortunately, many abstracts disagree with the text of the article (6). The format required for structured abstracts differs from journal to journal, and some journals use more than one structure; authors should make it a point prepare their abstracts in the format specified by the journal they have chosen.

Some journals request that, following the abstract, authors provide, and identify as such, 3 to 10 key words or short phrases that capture the main topics of the

article. These will assist indexers in cross-indexing the article and may be published with the abstract. Terms from the Medical Subject Headings (MeSH) list of Index Medicus should be used; if suitable MeSH terms are not yet available for recently introduced terms, present terms may be used.

Introduction

Provide a context or background for the study (i.e., the nature of the problem and its significance). State the specific purpose or research objective of, or hypothesis tested by, the study or observation; the research objective is often more sharply focused when stated as a question. Both the main and secondary objectives should be made clear, and any pre-specified subgroup analyses should be described. Give only strictly pertinent references and do not include data or conclusions from the work being reported.

Methods

The Methods section should include only information that was available at the time the plan or protocol for the study was written; all information obtained during the conduct of the study belongs in the Results section.

Selection and Description of Participants. Describe your selection of the observational or experimental participants (patients or laboratory animals, including controls) clearly, including eligibility and exclusion criteria and a description of the source population. Because the relevance of such variables as age and sex to the object of research is not always clear, authors should explain their use when they are included in a study report; for example, authors should explain why only subjects of certain ages were included or why women were excluded. The guiding principle should be clarity about how and why a

study was done in a particular way. When authors use variables such as race or ethnicity, they should define how they measured the variables and justify their relevance.

Technical information. Identify the methods, apparatus (give the manufacturer's name and address in parentheses), and procedures in sufficient detail to allow other workers to reproduce the results. Give references to established methods, including statistical methods (see below); provide references and brief descriptions for methods that have been published but are not well known; describe new or substantially modified methods, give reasons for using them, and evaluate their limitations. Identify precisely all drugs and chemicals used, including generic name(s), dose(s), and route(s) of administration.

Authors submitting review manuscripts should include a section describing the methods used for locating, selecting, extracting, and synthesizing data. These methods should also be summarized in the abstract.

Statistics. Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as the use of P values, which fails to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used.

Results

Present your results in logical sequence in the text, tables, and illustrations, giving the main or most important findings first. Do not repeat in the text all the data in the tables or illustrations; emphasize or summarize only important observations. Extra or supplementary materials and technical detail can be placed in an appendix where it will be accessible but will not interrupt the flow of the text; alternatively, it can be published only in the electronic version of the journal.

When data are summarized in the Results section, give numeric results not only as derivatives (for example, percentages) but also as the absolute numbers from which the derivatives were calculated, and specify the statistical methods used to analyze them. Restrict tables and figures to those needed to explain the argument of the paper and to assess its support. Use graphs as an alternative to tables with many entries; do not duplicate data in graphs and tables. Avoid non-technical uses of technical terms in statistics, such as “random” (which implies a randomizing device), “normal,” “significant,” “correlations,” and “sample.”

Where scientifically appropriate, analyses of the data by variables such as age and sex should be included.

Discussion

Emphasize the new and important aspects of the study and the conclusions that follow from them. Do not repeat in detail data or other material given in the Introduction or the Results section. For experimental studies it is useful to begin

the discussion by summarizing briefly the main findings, then explore possible mechanisms or explanations for these findings, compare and contrast the results with other relevant studies, state the limitations of the study, and explore the implications of the findings for future research and for clinical practice.

Link the conclusions with the goals of the study but avoid unqualified statements and conclusions not adequately supported by the data. In particular, authors should avoid making statements on economic benefits and costs unless their manuscript includes the appropriate economic data and analyses. Avoid claiming priority and alluding to work that has not been completed. State new hypotheses when warranted, but clearly label them as such.

References

General Considerations Related to References. Although references to review articles can be an efficient way of guiding readers to a body of literature, review articles do not always reflect original work accurately. Readers should therefore be provided with direct references to original research sources whenever possible. On the other hand, extensive lists of references to original work on a topic can use excessive space on the printed page. Small numbers of references to key original papers will often serve as well as more exhaustive lists, particularly since references can now be added to the electronic version of published papers, and since electronic literature searching allows readers to retrieve published literature efficiently.

Avoid using abstracts as references. References to papers accepted but not yet published should be designated as “in press” or “forthcoming”; authors should

obtain written permission to cite such papers as well as verification that they have been accepted for publication. Information from manuscripts submitted but not accepted should be cited in the text as “unpublished observations” with written permission from the source.

Avoid citing a “personal communication” unless it provides essential information not available from a public source, in which case the name of the person and date of communication should be cited in parentheses in the text. For scientific articles, authors should obtain written permission and confirmation of accuracy from the source of a personal communication.

Some journals check the accuracy of all reference citations, but not all journals do so, and citation errors sometimes appear in the published version of articles. To minimize such errors, authors should therefore verify references against the original documents. Authors are responsible for checking that none of the references cite retracted articles except in the context of referring to the retraction. For articles published in journals indexed in MEDLINE, the ICMJE considers PubMed the authoritative source for information about retractions. Authors can identify retracted articles in MEDLINE by using the following search term, where pt in square brackets stands for publication type: Retracted publication [pt] in PubMed.

Reference Style and Format. The Uniform Requirements style is based largely on an ANSI standard style adapted by the National Library of Medicine (NLM) for its databases. Authors should consult National Library of Medicine’s Citing Medicine for information on NLM’s recommended citation formats for a variety of reference types.

References should be numbered consecutively in the order in which they are first mentioned in the text. Identify references in text, tables, and legends by Arabic numerals in parentheses. References cited only in tables or figure legends should be numbered in accordance with the sequence established by the first identification in the text of the particular table or figure. The titles of journals should be abbreviated according to the style used in Index Medicus. Consult the list of Journals Indexed for MEDLINE, published annually as a separate publication by the National Library of Medicine. The list can also be obtained through the Library's web site. Journals vary on whether they ask authors to cite electronic references within parentheses in the text or in numbered references following the text. Authors should consult with the journal that they plan to submit their work to.

Tables

Tables capture information concisely, and display it efficiently; they also provide information at any desired level of detail and precision. Including data in tables rather than text frequently makes it possible to reduce the length of the text.

Type or print each table with double spacing on a separate sheet of paper. Number tables consecutively in the order of their first citation in the text and supply a brief title for each. Do not use internal horizontal or vertical lines. Give each column a short or abbreviated heading. Authors should place explanatory matter in footnotes, not in the heading. Explain in footnotes all non-standard abbreviations. For footnotes use the following symbols, in sequence:

*,†,‡,§,||,¶,**,††,‡‡

Identify statistical measures of variations, such as standard deviation and standard error of the mean. Be sure that each table is cited in the text. If you use data from another published or unpublished source, obtain permission and acknowledge them fully.

Additional tables containing backup data too extensive to publish in print may be appropriate for publication in the electronic version of the journal, deposited with an archival service, or made available to readers directly by the authors. In that event an appropriate statement will be added to the text. Submit such tables for consideration with the paper so that they will be available to the peer reviewers.

Reference

International Committee of Medical Journal Editors, 2007. *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication.* Retrieved September 25, 2008, from www.icmje.org

Section 5

Ethics approval letter

NHS

National Research Ethics Service

Nottingham Research Ethics Committee 1
1 Standard Court
Park Row
Nottingham
NG1 6GN

Telephone: 0115 9123344 ext 39368
Facsimile: 0115 9123300

27 November 2007

Miss Katie Morris
Trainee Clinical Psychologist
Trent Doctorate in Clinical Psychology
University of Lincoln
Health Life and Social Sciences
Court 11 Satellite Building 8
Brayford Pool
Lincoln LN6 7TS

Dear Miss Morris

Full title of study: **The validity of the Addenbrooke's Cognitive Examination Revised (ACE-R) as a screening measure for cognitive impairment in stroke**

REC reference number: **07/H0403/164**

The Research Ethics Committee reviewed the above application at the meeting held on 13 November 2007. Thank you for attending to discuss the study.

Ethical opinion

Discussion and clarification from you and your Supervisor:

The Committee agreed that there were no major ethical issues with this study.

- The Committee were unclear as to what method of analysis will be used i.e. what statistical analysis will be undertaken to see that the ACE-R is the superior test. Professor Lincoln confirmed that the ACE-R is more sensitive than other statistical analysis, it has been quantified and the ultimate cut-off point can also be identified

- The Committee were unsure as to whether junior doctors are happy to use ACE-R. You confirmed that they were happy to use it, as the mini-mental score is embedded in the ACE-R

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

The Committee agreed that all sites in this study should be exempt from site-specific assessment (SSA). There is no need to submit the Site-Specific Information Form to any Research Ethics Committee. The favourable opinion for the study applies to all sites involved in the research.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Application		06 August 2007
Investigator CV		06 August 2007
Protocol	1	06 August 2007
Participant Information Sheet: B	1	06 August 2007
Participant Information Sheet: A	1	06 August 2007
Participant Consent Form: (B)	1	06 August 2007
Participant Consent Form: (A)	1	06 August 2007
Participant Consent Form: Assent Form (B)	1	06 August 2007
Participant Consent Form: Assent Form (A)	1	06 August 2007
Mental Capacity Act supplementary information		
Neuropsychological Tests - Record Forms		
Addenbrooke's Cognitive Examination - ACE-R		
Patient Record Form		
CV - Key Investigator	2	01 August 2007
CV - Key Investigator	1	01 January 2006
Behavioural inattention test - scoring sheet		
The Hayling and Brixton Tests - Scoring Sheet		
FAS-Test Score Sheet		
Rey Complex Figure Test and Recognition Trial Test Booklet		
Sheffield Screening Test for Acquired Language Disorders		
Wechsler Memory Sacle - Third Edition		

R&D approval

You should arrange for the R&D office at all relevant NHS care organisations to be notified that the research will be taking place, and provide a copy of the REC application, the protocol and this letter.

All researchers and research collaborators who will be participating in the research at a NHS site must obtain final approval from the R&D office before commencing any research procedures.

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

Here you will find links to the following

a) Providing feedback. You are invited to give your view of the service that you have received from the National Research Ethics Service on the application procedure. If you wish to make your views known please use the feedback form available on the website.
b) Progress Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
c) Safety Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
d) Amendments. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
e) End of Study/Project. Please refer to the attached Standard conditions of approval by Research Ethics Committees.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nationalres.org.uk .

07/H0403/164	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

P. Wheat

PP **Dr K Pointon**
Chair

Email: trish.wheat@nottspct.nhs.uk

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
Standard approval conditions

Copy to: *Dr Mark Gresswell, University of Lincoln*
R&D office for NHS care organisation at lead site - NUH

Section 6

Appendices

Contents

Appendix 1

Additional background information

Appendix 1.1 Stroke and cognitive impairment

Stroke is defined as 'a focal (or at times global) neurological impairment of sudden onset, and lasting more than 24 hours (or leading to death), and of presumed vascular origin' (World Health Organisation, 1978). Stroke affects over 130,00 people in the UK every year and is the third most common cause of death. Stroke survivors may face a range of physical impairments following a stroke including, limb weakness, paralysis, muscle spacisity, pain, difficulty controlling bladder and bowels, problems swallowing and impaired vision. In addition, impairments of cognitive functioning are extremely common. Following the early stages post-stroke, impairments may improve over time as the brain recovers, or they may persist, causing long term disabilities (The Stroke Association, 2006).

'Cognitive functioning' is a term used to describe the abstract processes controlled by the brain such as language, perception, memory, reasoning, attention, psychomotor abilities (co-ordination of action), and 'executive' functions. (Executive functions is the name given to a number of higher order cognitive processes such as planning ability, abstract thinking, selective inhibition and cognitive flexibility. These are thought to be important in the execution of many everyday behaviours and in particular, the ability to respond adaptively in novel situations (Lezak, 2004).) Different cognitive functions are subserved by different parts of the cortex and subcortex.

As stroke can potentially cause damage to any part of the brain, a wide range of cognitive impairments are observed in stroke. Cognitive impairment is frequently focal in stroke, causing impairment in specific areas of functioning. However, depending on the location of stroke, or if there have been a series of strokes, there may be multiple impairments of cognition. Multiple impairment following stroke is known as vascular dementia (or 'multi-infarct dementia'). The DSM IV diagnostic criteria for dementia is impairment in memory, plus impairment in one of the following areas of cognition; language, psychomotor skills, perception, executive functioning (American Psychiatric Association, 1994). Thus, following stroke, patients with cognitive impairment can be classified into two groups; vascular cognitive impairment not dementia (VCIND) and vascular dementia (VaD).

The reported prevalence of cognitive impairment in stroke has differed, and studies have varied according to sample size, time since stroke, and the tests used to assess cognitive impairment. Nys et al. (2005a) report prevalence rates using a sample of 34 patients, in the early stages post-stroke (average 6.5 days). A strength of this study, above others, is the broad range of neuropsychological tests used to determine impairment. Cognitive domains assessed were abstract reasoning, verbal memory, executive functioning, visual perception and construction, visual memory and language. In this study, the prevalence of cognitive impairment in one or more areas of cognitive functioning was found to be 70%. In a later study, Nys and colleagues reported that the risk of cognitive impairment varies according to location of stroke. In patients with cortical stroke, the risk of cognitive impairment is greatest at 74%. Patients

with subcortical and infratentorial stroke have a risk of cognitive impairment of 46% and 43% respectively (Nys et al., 2007b).

Reports vary as to the most commonly impaired areas of cognition in the acute phases post-stroke, and findings are subject to the different methods used between studies. For example, in Nys et al. (2005a), the most common cognitive impairments were reported to be disorders of abstract reasoning (47%), followed by impairments in executive functioning (32%), language (26%), visual perception and construction (21%), verbal memory (15%) and visual memory (12%). However, Lesniak, Bak, Czepiel, Seniow and Czlonkowska (2008) report very different findings. This study measured patients at a similar time post-stroke to Nys et al. (2005a) (during the second week), but used a larger sample (200 patients), and a different battery of tests. They found that the most common cognitive impairment was attentional functioning (48.5%), followed by language (27%), short term memory (24.5%), and executive functioning (18.5%). A review of the literature in this area cites executive functioning as the most commonly affected area (Engstad, Viitanen & Almkvist, 2007). However, this review includes studies which either use poorly validated measures of cognitive impairment and/or examine a smaller range of cognitions. Therefore, overall, we cannot make any conclusions as to the most commonly affected areas of cognition post stroke.

Following the acute stages of stroke, many patients with cognitive impairment recover, whilst in some patients deficits remain stable, and in some patients deficits deteriorate further. In a longitudinal study, Appelros and Andersson

(2006) found 32% of patients with cognitive impairment had deteriorated further after one year, 13% had remained unchanged, and 55% had improved. Three months post-stroke, Madureira, Guerreiro and Ferro (2002) found that 55% of all stroke patients were impaired on at least one domain of cognitive functioning, as measured on a comprehensive neuropsychological battery of tests. Of these impaired patients, 49% had VCIND and 6% had VaD. Patel, Coshall, Rudd and Wolfe (2003) assessed patients at three months, one, two and three years post stroke and found rates of cognitive impairment to be 39%, 35%, 30% and 32% respectively. It is possible, however, that the rates found by Patel et al. may be underestimated due to use of a limited degree of cognitive testing in this study. Prognosis has been found to be most favourable for impairments in perception/construction and visual memory, and cognitive recovery has been found to be associated with age, pre-existing ability, lesion volume, lesion location and diabetes (Nys et al, 2005b).

The presence of cognitive impairment has a significant impact on an individual's life post-stroke. Impairments in cognitive functioning post-stroke are strongly associated with a person's ability to carry out tasks of daily living (ADLs) and an ability to live independently, regardless of the person's physical ability (Claesson, Linden, Skoog & Blomstrand, 2005; Nys et al., 2005c).

Appendix 1.2 Principles of cognitive screening in acute stroke

Thorough neuropsychological assessment of cognitive impairment is time consuming, costly and can prove too tiring or too difficult for unwell or very impaired patients. In the early stages post stroke, rapid improvement also means that any assessment is invalidated after a relatively short period of time. Thus, in acute stroke, it is more appropriate, in the first instance, to use a brief cognitive screening measure to identify those people with cognitive problems and to highlight areas for more in depth assessment. The screening measure should be quick and easy to administer, suitable for bedside (hospital) administration, and suitable to be given by staff without specialised training in assessment. Cognitive screening for all stroke patients is recommended by the National Clinical Guidelines for Stroke (Royal College of Physicians of London, 2004). Whilst the guidelines do not specify which screening instrument should be used, it is advised that clinicians select a measure that has suitable validity (appropriateness for the purpose).

A test used for diagnosis or screening should have good diagnostic validity, i.e. ability to accurately detect presence of impairment in a given population. This requires high levels of *sensitivity* and high levels of *specificity* (Loong, 2003). Sensitivity refers to how good a test is at correctly identifying those people in a population who have the impairment or disease (i.e. proportion of true positives identified by the test). Specificity refers to how good a test is at correctly identifying those people who do not have the impairment or disease (i.e. proportion of true negatives identified by the test). Thus a sensitivity of 80%

means the test will correctly detect the impairment/disease in 80% of those people affected, and it will miss 20% of people affected. A specificity of 90% means the test will correctly classify 90% of unimpaired people as being unimpaired, and 10% of unimpaired people will be misclassified as impaired when they are not. Sensitivity and specificity are statistically linked and trade-off each other. Therefore, increasing the sensitivity of a test (i.e. by lowering cut off scores) decreases specificity. Diagnostic measures, as opposed to screening measures, should generally have a sensitivity of greater than 80% and a specificity of greater than 60% (Lincoln, Nicholl & Flannaghan, 2003). In a screening test, which pre-empts full testing, it may be better to increase sensitivity to 90%, in order to detect most of those people with problems. However, it is important that specificity is not compromised too much, as low specificity will lead to full testing on lots of people who do not have impairment. This would mean the screening test would have little purpose, and it would be as well to do the full testing on everyone (Lincoln et al., 2003). The 'Area Under the Curve' (AUC) statistic is sometimes used in studies to summarise the sensitivity and specificity of a test. This figure refers to the overall ability of a test to correctly classify those people with or without the disease, specifically, the probability that the result for a randomly chosen positive case will exceed the result for a randomly chosen negative case). AUC=1 is a perfect test (meaning 100% sensitivity and 100% specificity), .90-1 is excellent, .80-.90 is good, .70-.80 is fair, .60-.70 is poor and AUC 0.5 is worthless (suggesting discriminative ability is the same as chance).

When considering the diagnostic validity of a test, it is also useful to consider *positive and negative predictive power statistics* (Loong, 2003). A positive predictive value takes all positive test results (true positives and false positives), and describes the chance of a positive test result being correct. Negative predictive values take all negative results, and describe the chances of a negative result being correct. Predictive power statistics take into account not only the diagnostic accuracy (sensitivity and specificity) of the test, but also the natural prevalence ('base rate'*)* of impairment or disease in a population. This is important because base rates can affect the likelihood of a true positive or true negative result, over and above the accuracy of the particular test being used. For example, in a population where there is high prevalence of impairment, a positive test result is more likely to be correct compared to a positive test result using the same test in a population with a low prevalence rate, regardless of the accuracy of the test. In an acute stroke population, high base rates of cognitive impairment of around 70% (Nys et al. (2005a) will naturally contribute towards higher positive predictive power, and lower negative predictive power. Considering base rates alone, we can be 70% confident that a person will have cognitive impairment, but only 30% confident a person doesn't have cognitive impairment. It might therefore be assumed that there is little clinical value to using a screening measure in stroke which has a PPV of less than 70% or a NPV of less than 30%.

To summarise, a screening test should have good positive and negative predictive values as well as high sensitivity and specificity. A shortcoming of

many existing validity studies is a failure to report positive and negative predictive values.

Appendix 1.3 Commonly used cognitive screening measures

1.3.1 The MMSE

The MMSE (Folstein, Folstein & McHugh, 1975) is currently the most widely used screening measure in acute stroke (Shulman, Hermann & Brodaty, 2006). The MMSE was designed to screen for cognitive impairment in a range of medical patients including dementia and psychiatric patients. Its purpose is to detect the presence and severity of cognitive impairment, and to map deteriorations and improvements over time. The test takes approximately five minutes to administer and comprises of 11 items assessing, orientation, verbal memory recall, attention, language and perception/construction. The test produces a total score out of 30. Validity studies of the MMSE have identified two optimal cut-off scores for the detection of cognitive impairment (27 and 24). It is recommended clinicians use the higher or lower cut-off depending on the patient population and whether the clinician is more anxious to avoid false negatives or false positives. A lower cut-off of 24 is recommended for circumstances where it is necessary to ensure high diagnostic specificity, for example, when discriminating between dementia and other possible causes of cognitive impairment such as depression (e.g. Folstein et al., 1975; Yang, Hwang, Jen & Liu, 2000).

Although, the MMSE was designed to detect cognitive impairment in a range of medical conditions, by far its most common usage is for screening for dementia in community geriatric populations. A wealth of studies have examined the validity of the MMSE in this field, most studies using populations of Alzheimer's type dementia (DAT) or mixed dementia populations. On the whole, studies

have shown the MMSE to have high levels of diagnostic validity for detecting dementia (e.g. Tombaugh & McIntyre, 1992). For example, Helkala et al. (2002), found the MMSE (cut-off score of 24) to have sensitivity of 82% and specificity of 64% (compared with diagnosis achieved through medical history, neurological exam, neuropsychological tests, MRI, CSF, ECG, chest exams and blood tests).

Despite this, the MMSE has been found to have several limitations, including inaccurate diagnosis of early/mild impairment, and variation of normal scores according to age, IQ and cultural background (Tombaugh & McIntyre, 1992; Nadler, Richardson & Malloy, 1994; Bieliauskas, Depp, Kauszler, Steinberg & Lacy, 2000; Starr & Lonie, 2007; Trenkle, Shankle & Azen, 2007). The MMSE has also been shown to have limitations for the accurate detection of cognitive impairment in dementias and other organic disorders with subcortical impairment, frontal lobe impairment, or right hemisphere dysfunction (Naugle & Kawczak, 1989; Dick et al., 1994; Grace et al., 1995; Darvesh & Freedman, 1996). In acute stroke populations, which tend to characterise the above patterns of impairment, diagnostic validity is also poor (Blake, McKinney, Treece, Lee & Lincoln, 2002; Nys et al., 2005a).

Nys et al. (2005a) found that at a cut-off score of 24, the MMSE had fair specificity (70%), but extremely poor sensitivity (34%) for the detection of cognitive impairment in stroke. No alternative cut off scores on the MMSE could yield both a sensitivity of greater than 80% and a specificity of greater than 60%. Overall, when differentiating patients with cognitive impairment and

those without, the MMSE performed statistically no better than chance, AUC of 0.67 (p=0.13). Blake et al. (2002) found the MMSE, at a cut-off of 24, had better levels of sensitivity and specificity in acute stroke (62% and 88% respectively), although this still represents inadequate levels of sensitivity. However, Nys et al. (2005a) criticised the methodology of this study, highlighting the large individual differences in intervals between the administration of the MMSE and the neuropsychological examination used to determine the presence of impairment (up to 3 months). Given that recovery is greatest in the first three months post-stroke, Nys et al. (2005a) suggested that patients may have performed much better on the neuropsychological examination, thus over estimating the sensitivity of the MMSE in this study. In addition, as Blake et al. (2002) noted, patients did not complete a standard selection of tests as part of their neuropsychological examination, in order to limit time spent on testing. This means that there is a possibility that patients may have performed, overall, better on the neuropsychological examination relative to their true ability, further overestimating the sensitivity of the MMSE.

The poor validity of the MMSE in acute stroke, compared to in dementia populations (consisting mainly of DAT), is probably explained by inadequate assessment of different domains of cognitive functioning. Compared to DAT, there is a more heterogeneous range of impairments seen in stroke, due to both cortical and subcortical impairments, and the existence of focal as well as global impairments. The MMSE assesses only a narrow range of cognition (verbal memory recall, attention, language and perception/construction), omitting any measure of executive functioning and assessing very few performance abilities.

Nys et al. (2005a) found that, compared to detailed neuropsychological assessments, the MMSE was most likely to miss impairments in reasoning, executive functioning, and visual perception/construction.

Besides having poor validity, another limitation of the MMSE in stroke is that it only provides normative data for the overall score, and it does not have norms for individual subscales. Therefore, it only has the capacity to identify the presence or absence of cognitive impairment per se, and it is not possible to make any quantitative conclusions about severity of impairment in specific areas of cognitive functioning (e.g. memory). It is also not reliable to make *qualitative* conclusions from performance on subtests as the subtests of the MMSE have been found to have poor domain specificity (Giordani, Boivin, Hall & Foster, 1990; Tierney, Szalai, Snow, Fisher & Dunn, 1997). A measure of this kind has disadvantages in a stroke population. Here, it is less relevant to know whether someone has cognitive impairment per se (it already likely they do due to high base rates). It is more useful to know what specific problems exist, in order to inform rehabilitation and provision of care. Domain specific cognitive abilities in acute stroke also have prognostic value in terms of long term cognitive and functional outcomes (Nys et al., 2005c).

To summarise, the MMSE was designed for the detection of cognitive impairment in a wide range of medical conditions. It is most commonly used in community geriatric patients for the detection of dementia (mainly DAT), where its diagnostic validity has been found to be good. However the MMSE has limited validity for the detection of cognitive impairment in organic conditions,

such as stroke, where impairment is more heterogeneous than in DAT. This is probably due to the assessment of a restricted range of cognitive domains. Furthermore, the MMSE has restricted utility in stroke because it is unable to give information about the nature of cognitive impairment in individuals.

1.3.2 The Cambridge Cognitive Examination (CAMCOG).

The CAMCOG forms part of the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) (Roth et al., 1986). This screening test was designed specifically to assess the range of cognitive functions required for a diagnosis of dementia. It is a more comprehensive screening measure compared to the MMSE, with more items, covering a larger number of cognitive domains (orientation, language, memory, attention, praxis, calculation, abstract thinking and perception). The CAMCOG comprises of all items of the MMSE plus additional questions, yielding a MMSE score out of 30 and a CAMCOG score out of 106. Clinicians can also calculate scores for several individual cognitive domains allowing analysis of relative impairment in specific areas.

In the original validation study (Roth et al., 1986), the authors found that a CAMCOG total cut-off score of 79/80 was optimal for the discrimination of demented patients (DAT, VaD or other dementias) from non-demented patients. This CAMCOG cut-off score gave a sensitivity of 92% and specificity of 96%. In comparison, the optimal MMSE cut-off score in this study was found to be 23/24 which gave similar levels of sensitivity (94%) but poorer specificity (85%). Unlike the MMSE, the CAMCOG did not show ceiling effects and was sensitive to milder degrees of cognitive impairment. Similar results have been reported in

later studies (Blessed, Black, Butler & Kay, 1991; Huppert, Brayne, Gill, Paykel & Beardsall, 1995).

De-Koning et al. (1998) have also shown the CAMCOG to have excellent validity for the detection of dementia syndrome in a post-stroke population. The overall diagnostic accuracy of the CAMCOG in this study for detecting dementia was found to be 95% (AUC = 0.95), which was somewhat superior to the MMSE (AUC=0.90). (The authors did not report sensitivity and specificity values for specific cut-off scores.) Despite its promising psychometric properties for use in stroke populations, a major disadvantage of the CAMCOG for use in stroke is the time taken to administer the test (approximately 25 minutes).

To summarise, the CAMCOG offers a more comprehensive assessment of cognitive functioning compared to the MMSE, and it has been shown to have superior diagnostic validity for the detection of dementia. However, the trade-off for this is longer administration time.

1.3.3 The Rotterdam-Cambridge Cognitive Examination (R- CAMCOG).

De-Koning, Dippel, van-Kaooten and Koudastall (2000) aimed to address the shortcomings of the CAMCOG by developing a shortened version, the 'Rotterdam-CAMCOG' (R-CAMCOG) for use in stroke. This new measure takes just 10 minutes to administer, with the reduction in length achieved by removing items with floor and ceiling effects and items with poor diagnostic validity. A preliminary analysis of the validity of the R-CAMCOG on the dataset

used to develop the test suggested the R-CAMCOG to have equal levels of overall diagnostic accuracy as the CAMCOG (AUC= 0.95 for both tests) (de-Koning et al., 2000). An optimal cut-off score of 33/34 was recommended for the detection of post-stroke dementia.

In a later, separate, validation study, analysis confirmed the R-CAMCOG to have good validity for the detection of dementia post-stroke at the recommended cut-off; sensitivity= 66% and specificity= 94% (de-Koning, van-Kooten, Koudstaal & Dippel, 2005). A short-coming of this study, however, was the failure to compare this with validity of the MMSE within the same sample. Instead, the authors compared figures to those found in their earlier study (de Koning et al., 1998). This indicates the R-CAMCOG to have better specificity than the MMSE (94% verses 84%), but slightly worse sensitivity (66% verses 69%). Unfortunately, the authors do not make a comparison with the validity of the CAMCOG found in the earlier study, and it is not possible for the reader to do so independently as different statistics are reported in the two studies.

Despite minor methodological shortcomings, the findings of de-Koning and colleagues are encouraging, suggesting that the R-CAMCOG may have good diagnostic accuracy for the detection of dementia post-stroke. However, it is important to note that de-Koning et al.'s studies only examined validity for detecting *dementia* post-stroke (VaD), and they did not look at ability to detect VCIND. Thus, we do not know how good the CAMCOG/R-CAMCOG would be at detecting impairments in an inclusive stroke population.

Given that there is no measure of executive functioning in these tests, it seems reasonable to predict that the CAMCOG/R-CAMCOG may be less accurate at identifying cognitive impairment in a stroke population which includes patients with VCIND. Like the MMSE, the CAMCOG/R-CAMCOG may misdiagnose patients whose impairments are characterised by problems in executive functioning (Nys, 2005a). This probably limits the potential of the CAMCOG as a screening test for all stroke patients. A further limitation of this test as a screen in stroke is that it requires specialised test equipment, making it unsuitable for bedside administration.

To summarise, the R-CAMCOG is a shortened version of the CAMCOG, designed specially for the detection of dementia post-stroke. Validity studies have suffered from some methodological short-comings but indicate comparable levels of validity compared to both the CAMCOG and the MMSE. However, its utility for the detection of cognitive impairment in inclusive acute stroke populations remains unclear due to lack of studies, and its potential for this purpose may be limited due to the omission of a measure of executive functioning. Furthermore, the test is likely to be impractical as a screening measure, requiring specialised test equipment.

1.3.4 The Addenbrooke's Cognitive Examination (ACE).

The ACE (Mathuranath, Nestor, Berrios, Rakowicz & Hodges, 2000) was developed in order to overcome the limitations of the MMSE and the CAMCOG, that is, to provide a measure which is sensitive to a larger range of cognitive impairment, a measure sensitive to mild degrees of impairment, and a measure

suitable for bedside administration. In addition, the ACE was designed to detect and differentiate Alzheimer's disease (DAT) and frontotemporal dementia (FTD). The ACE comprises the MMSE plus additional items which expand on memory, language and visuospatial components. It also includes a measure of executive functioning (verbal fluency) which the MMSE omits. Six cognitive domains are assessed in total; Attention, Orientation, Memory, Fluency, Visuospatial and Language. The memory section evaluates episodic memory (recall of three items from the MMSE plus name and address learning) and semantic memory. The language component assesses naming (of line drawings), comprehension, repeating spoken words, reading words and writing a sentence. The visuospatial component consists of copying overlapping pentagons (from the MMSE), copying a wire cube and drawing a clock face. The fluency section comprises of a letter fluency and category fluency test. The ACE gives a total score out of 100, as well as a MMSE score, out of 30, making it easy to compare performance to previous MMSE scores documented by other clinicians. In addition to total scores, the ACE also gives scores on the six subscales. Maximum scores are; orientation (10 points), attention (eight points), memory (35 points), verbal fluency (14 points), language (28 points), and visuospatial ability (five points). The ACE is a test requiring only pen and paper and takes between 15 and 20 minutes to administer.

In the original validation study (Mathuranath et al., 2000), normative data was provided for the ACE from a sample of 127 patients from non-dementia clinics (orthopaedic and gynaecological clinics). 139 patients from a dementia clinic were also recruited and the ACE was examined for its ability to accurately discriminate the presence of dementia in the whole sample. Validity was

compared to the MMSE and reliability of the ACE was also examined. Diagnosis of dementia had been determined by CT or MRI scan, laboratory evaluation, and neuropsychiatric/neuropsychological test batteries. Dementia diagnoses were of mixed type including; AD, VaD, FTD dementia with Lewy bodies, and corticobasal degeneration.

The validity of the ACE for detecting dementia was considered at a range of cut-off scores. The first cut-off score examined was a score of 88, which represents the mean score of the control group plus two standard deviations. At this cut-off, the ACE had a sensitivity of 93% and a specificity of 71% for detecting dementia. Analysis determined an alternative cut-off of 83 which enhanced specificity (96%), whilst compromising sensitivity (82%). Predictive values at a range of reported prevalence rates were reasonably high for both cut-off scores. In comparison, the optimal cut-off score on the MMSE was found to be 27, yielding a sensitivity of 74% and a specificity of 96%. A cut-off of 24 on the MMSE yielded comparable levels of specificity (96%), whilst greatly reducing sensitivity (52%). The reliability of the ACE was measured in terms of internal consistency. Cronbach's alpha for the ACE was very good (0.78).

Mathuranath et al. (2000) concluded that the ACE had better sensitivity, but equal levels of specificity compared to the MMSE for detecting dementia. The ACE was particularly superior compared to the MMSE when using the lower cut-offs (83 and 24 respectively). The authors recommended use of an ACE cut-off score of 88 in populations where it is more important to reduce false negatives (i.e. clinical settings, particularly where the base rate of dementia is

high), and use of a cut-off score of 83 where it is more important to reduce false positives (e.g. screening for a therapeutic trial or in clinical settings where the base rate for dementia is low).

On the basis of further sub-group analysis, the authors also concluded that the ACE maintained good sensitivity for diagnosing different severity and types of dementia. The ACE proved particularly advantageous over the MMSE at detecting FTD patients, nearly doubling the detection rate. Furthermore, the calculation of a 'VLOM' ratio (ratio of scores on Fluency plus Language to Orientation plus name and address delayed recall memory) could accurately discriminate FTD from non-FTD patients.

Further to the original validation study, Dudas (2005) compared the ACE scores of patients who had cognitive impairment due to dementia and the ACE scores of patients who had cognitive impairment due to affective disorder. The results showed that the total ACE scores of the dementia groups were significantly lower (<88) compared to total ACE scores of the affective disorder group. In addition, the patient groups differed in their subscales profiles; patients with affective disorder presented with mild impairment on ACE total scores plus low memory and fluency scores, whilst patients with dementia were characterised by low total ACE scores. Thus adding to the findings of the original study, this study indicates that, not only can the ACE can differentiate dementia patients from controls, but it can also differentiate dementia patients from other patients with cognitive impairment.

One limitation of the index study (Mathuranath et al., 2000) was its reliance on cross-sectional data in a patient group where a dementia diagnosis is only ever 'probable' and diagnosis, as determined by the gold standard clinical assessment, occasionally changes over time. Larner (2005, 2006) addressed this limitation by undertaking a longitudinal audit in a Cognitive Function Clinic, publishing two consecutive papers, examining different properties of the ACE.

Larner's first paper (Larner, 2005) reported the ability of the ACE to discriminate the presence of dementia in 154 consecutive new patients referred to the clinic. Diagnosis of dementia was confirmed with a standard patient and informant semi-structured interview, formal neuropsychological assessment and structural brain imaging (CT or MRI scan). Patients had mixed dementia diagnoses including DAT, FTD, VaD, VCIND, dementia with Lewy bodies and alcohol-related dementia. This study very much replicated the index study in design, with the key difference being that patients had a minimum of six months follow-up to confirm or establish diagnosis. At follow-up, 80 patients were found to have dementia, and 78 patients were not demented. For the diagnosis of dementia, an ACE cut-off score of 88 in this study gave a sensitivity of 97% and specificity of 47%. A cut-off score of 83 gave a sensitivity of 92% and specificity of 62%. A MMSE cut-off score of 27 gave a sensitivity of 91% and a specificity of 70%, and a MMSE of 24 gave a sensitivity of 73% and a specificity of 86%. Thus the ACE offered greater sensitivity but lesser specificity than the MMSE. In comparison to the index study (Mathuranath et al., 2000), the results are similar in terms of sensitivity but worse in terms of specificity. The author pointed out that the poor specificity may be due to the unselected nature of the

patient cohort in their study. All patients had cognitive dysfunction of some kind (hence being referred to a cognitive dysfunction clinic) and thus many would fail the screens for reasons other than dementia e.g. depression.

In a second paper, Larner (Larner, 2006) reported the ability of the ACE for tracking cognitive change over time in 23 patients whose diagnoses had originally been uncertain. Patients were administered the ACE between 2-5 times over a 7-36 month period. As previously, an initial clinical diagnosis was based on standard patient and informant semi-structured interview, formal neuropsychological assessment and structural brain imaging (CT or MRI scan). Clinical diagnosis at the end of the follow-up period was based on care-giver report and clinical judgement. The author compared the diagnostic accuracy of the first ACE scores to the diagnostic accuracy of the final ACE scores. Comparing the first and last assessments, sensitivity and specificity rates remained comparable and positive predictive values improved. Larner (2006), whilst acknowledging the small sample size in this cohort, suggests that the ACE is responsive to longitudinal change in dementia.

To summarise, the ACE was developed as a bedside screening measure which would overcome the limitations of the MMSE, specifically, assessing a wider range of cognitive domains and allowing more in-depth analysis by the provision of subscale scores. Validity studies have indicated that the ACE has superior sensitivity compared to the MMSE in the detection of dementia, but comparable or marginally worse levels of specificity. Analysis of the ACE subscales enable the accurate discrimination of different types of cognitive impairment, and the ACE is sensitive to longitudinal change in cognition.

1.3.5 The Addenbrooke's Cognitive Examination-Revised (ACE-R).

Since the ACE, a revised version of the ACE, the ACE-R, has been published, which has now superseded the use of the ACE in clinical practice (Mioshi, Dawson, Mitchell, Arnold & Hodges, 2006). In developing the ACE-R, the ACE underwent several design modifications in order to make it easier and quicker to administer, to increase the test sensitivity and to facilitate cross cultural usage and translation. In addition, three different versions of the ACE-R were designed to allow for repeat testing. Modifications to the content of the test included major changes to the memory component, reducing its weight in the overall score and allowing other domains to have a more balanced contribution. The fluency scoring system was adjusted, and some changes to the language section were made, including increasing the difficulty of the naming task. The visuospatial section was expanded to include counting of dot arrays and identification of fragmented letters. There were no changes to the attention and orientation sections. The maximum score on the ACE-R remains 100, but the subscales are five instead of six, with attention and orientation being combined. They are weighted as follows; attention and orientation (18 points), memory (26 points), fluency (14 points), language (26 points) and visuospatial (16 points). The ACE-R takes marginally less time than the ACE to administer (between 12-20 minutes).

In Mioshi et al., 2006 the methodology for the validation of the ACE-R was similar to that of the original ACE study (Mathuranath et al., 2000). Control participants were used to create new normative data for the ACE-R total score,

and the diagnostic validity of the ACE-R for detecting dementia was determined at a variety of cut-offs. Unlike in the original study, normative data was also provided for the test's subscales to allow direct comparison of scores to normal controls' performance. A total of 241 subjects participated, completing both the ACE and the ACE-R. There were 142 patients with different diagnoses of dementia, 36 patients with mild cognitive impairment (MCI) and 63 controls. As before, diagnoses of dementia were determined by CT or MRI scan, laboratory evaluation, and neuropsychiatric/neuropsychological test batteries. Diagnosis of MCI was determined by performance on cognitive and functional tests and clinical interview.

In Mioshi et al. (2006) two possible ACE-R cut-off scores were identified in this study (88 and 82). At 88 the ACE-R gave excellent sensitivity (94%) and good specificity (89%) for the detection of dementia. A cut-off score of 82 reduced sensitivity (84%), whilst increasing specificity (100%). A cut-off of 82 gave perfect positive predictive value (100%) at all the prevalence rates investigated, whilst at a cut-off of 88, the ACE-R had lower positive predictive value, especially at low prevalence rates. Diagnostic validity of the ACE-R was not compared with that of the ACE or the MMSE in this study population, although given the methodology is similar to that in the original ACE study, one can assume a fair comparison. Considering the sensitivity and specificity rates reported in the early study, the ACE-R performed better than both the ACE and the MMSE at both high and low cut-offs, with the exception of the MMSE having better specificity at the higher cut-off of 27. The authors of this study did compare mean scores of participants on the ACE and the ACE-R, finding that there was indeed a significant difference in total ACE/ACE-R scores and on

most subscales. Thus, the authors concluded that the ACE-R resulted in better sensitivity and specificity and had striking levels of positive predictive value for detecting dementia at a range of prevalence rates.

Extended analysis in Mioshi et al. (2006) revealed significant differences in scores between dementia and MCI subjects, and between MCI and control subjects, suggesting the ACE-R is also sensitive to detecting MCI. The authors suggested, on the basis of clinical experience, a cut-off of 80 could be used for detecting MCI. In addition, analysis of the 'VLOM' ratio indicated that the ACE-R, like the ACE, could accurately discriminate FTD from non-FTD patients, although the authors noted specificity is much better than the sensitivity when used for this purpose.

Whilst Mioshi et al's finding were promising, they acknowledged that the study was developed within a very specialised population of patients (university clinic), and recommended that the ACE-R also be studied in community samples where disease characteristics and prevalence rates may be different. Larner (2007) conducted a pragmatic study of the ACE-R reflecting more day-to-day clinical practice. As in early studies by the same author, this study recruited patients from a cognitive function clinic, where there was no selection of patients by disease category, no application of exclusion criteria, and no control group of normal individuals- all patients had some form of cognitive dysfunction. The study examined the ability of the ACE-R to identify patients with a dementia (types included DAT, FTD, VaD and dementia with Lewy bodies), at the previously published cut-offs of 88 and 82. Compared to the

index study (Mioshi et al, (2006), the author found the ACE-R to have better sensitivity at both cut-offs of 88 and 82 (100% and 96% respectively). However, specificity rates were much poorer (48% and 72% respectively), as were positive predictive values (62% and 75% respectively). Given the different case mix of this sample compared to the index study (i.e. all patients had some form of cognitive impairment), the author felt justified to investigate a lower cut-off value (75). At this cut-off, excellent sensitivity was maintained (91%), specificity was greatly improved (91%) as was positive predictive value (89%). Thus the author concluded that the ACE-R had good diagnostic accuracy for detecting dementia in day-to-day clinical practice, but suggested lower cut-offs may be necessary depending on the setting and the case mix of referrals.

Whilst the vast majority of studies have examined the ACE and ACE-R for use in dementia populations, one recent study has examined the ability of the ACE-R to detect cognitive impairment in brain injured patients (Gaber, 2008). This study comprised of a sample of 36 brain injured patients, all with cognitive impairments severe enough to stop the patient returning to full time employment. The authors examined the ability of the ACE-R to correctly determine the presence of cognitive impairment in these patients at the two published cut-off scores of 88 and 82. Results showed the ACE-R to have better sensitivity than the MMSE at both the higher cut-off (ACE-R= 72%, MMSE= 36%), and at the lower cut-off score (ACE-R= 56%, MMSE= 11%). This study is limited, however, by the absence of a non-impaired control group, thereby making it impossible to consider specificity levels and rates of false positive diagnosis. Despite the possible limitations of this study, the fact that

the ACE-R seems to have, at least, good sensitivity in the field of brain injury is a good indication that the ACE-R may also perform well in stroke. This is because the cognitive presentation of these two groups of patients share more similarity than either do with dementia, with high rates of subcortical and frontal lobe impairment.

To summarise, the ACE-R was developed following modifications of the ACE aimed at decreasing administration time, increasing test sensitivity and facilitating cross-cultural usage. The ACE-R also offers the advantage over the ACE of having normative data for each subscale, allowing more accurate analysis of patterns of impairment. Validity studies have indicated that the ACE-R has increased sensitivity and specificity for the detection of dementia, compared to the ACE and the MMSE, although there has been some debate over the ideal cut-off score. One study has also given early indications that the ACE-R has utility for the detection of cognitive impairment in brain injury. The validity of the ACE/ACE-R has not yet been examined in stroke patients and this is the focus of the current study.

Appendix 1.4 The Sheffield Screening Test

The Sheffield Screening Test for Acquired Language Disorders (Syder, Body, Parker & Boddy., 1993) was used to screen for aphasia, which was one of the exclusion criteria for the study. The test comprises of two sections, addressing expressive and receptive language separately. Each section includes five subtests examining specific areas of linguistic processing. The receptive skills section examines; verbal comprehension of single words, comprehension of a sequential command, comprehension of a complex command, recognition of differences in meaning between words, and comprehension of narrative. The expressive skills section examines; word finding, abstract word finding, sequencing, ability to produce word definitions and verbal reasoning. Points are awarded for each correct answer and these are added to equal a total score on the test (max= 20). The authors of the test found age to be a significant independent variable in determining scores on the test, thus they recommended the use of a sliding cut-off score, on the basis of age. This attempts to eradicate the risk of false negatives in older people with non-pathological language deterioration. The cut-off scores for impairment, used in the current study, as shown in the test manual are <59 years = 17, 60-69 years = 16, and 70 + years = 15. The authors of the Sheffield Screening test report good concurrent and diagnostic validity and good inter-rater reliability. The normative sample was of an adequate size (n=112) and consisted of adults mostly over the age of 60, making it suitable for use with a stroke population, whom are mostly older adults. The Sheffield Screening Test is also suitable as it can be used by someone who is not a speech therapist and is a brief measure. There

have been subsequent studies that have included validation of the Sheffield Screening Test.

Appendix 1.5 The neuropsychological test battery

1.5.1 F,A,S verbal fluency

In the neuropsychological test battery, a *verbal fluency* test was used as a measure of executive functioning. Verbal fluency tests, originally developed by Berg (1948) are thought to measure cognitive flexibility (ability to selectively attend to relevant stimuli and filter out irrelevant stimuli). Verbal fluency tests have been shown to be highly sensitive to the presence of lesions in the frontal lobe (the area of the brain assumed to subserve executive functioning) (e.g. Henry & Crawford, 2004). They have also been found to be more sensitive to frontal lesions than other measures of cognitive flexibility, such as the Wisconsin Card Sorting test (Henry & Crawford, 2004). A huge number of versions of Verbal Fluency tests have been devised since the original paradigm was developed by Berg (1948). Tests commonly require subjects to generate words, at speed, beginning with specific letters of the alphabet, commonly F,A,S (phonemic fluency), or belonging to certain semantic categories (category fluency). The tests yield an overall score, equal to the total number of words produced for all three letters or categories. In the current study, phonemic fluency was assessed, rather than category fluency, as category fluency, whilst sensitive to frontal impairments, is also sensitive to temporal lobe impairments i.e. language dysfunction (Henry & Crawford, 2004). A large number of normative data sets exist for verbal fluency tests. In the current study, the norms produced by Tombaugh, Kozak & Rees (1999) were used. This gives normative data for the total number of words produced for the letters F,A,S, with 1 minute allowed for each letter). These norms are derived from a very large

sample (n=895), covering a wide age range (16-95 years). The data is also stratified by age and years of education, two variables commonly known to significantly predict verbal fluency. To ensure maximum validity of norms, the current study utilized the specific instructions used in the Tombaugh et al. (1999).

1.5.2 Hayling Sentence Completion

The Hayling Sentence Completion test from the Hayling and Brixton Tests (Burgess & Shallice, 1997) was used as another measure of executive functioning. This is thought to measure two abilities; response initiation and response suppression, both of which have been shown to be impaired in patients with frontal lobe lesions. The test consists of two sections. In both sections, the examiner reads aloud sentences which have a word missing at the end (e.g. "the old house will be torn…."). In section 1, subjects are required to complete the sentence sensibly, as quickly as possible, (e.g. "the old house will be torn *down*"). In Section 2, subjects are required to complete the sentence as quickly as possible with a word which does *not* fit at the end of sentence, a word that is unconnected in meaning (e.g. "the captain wanted to stay with the sinking *light bulb*"). The test yields three scores; the sum of latencies in section 1 (indication of response initiation), the sum of latencies on section 2 and the sum of errors on section 2 (indications of response suppression). In the original validation study by Burgess and Shallice (1997), the test was shown to have good diagnostic validity, accurately discriminating control subjects from patients with frontal lesions. Inter-rater reliability was shown to be very good, and split-half reliability was fair. However, the normative data provided in the original

study was not deemed suitable for use in the current study as it is based on a sample which included very few older adults (66-89 years n=19). Thus, a review of the literature was carried out to find an alternative normative data set. From this, the data provided in Bielak, Mansueti, Strauss and Dixon (2006) seemed the most suitable for the current study, consisting of a large sample (n= 457) of older adults (aged 53-90). Normative data is stratified by age in this study, as this was found to be a significant predictor of test performance.

1.5.3 Star Cancellation

The Star Cancellation test from the Behavioural Inattention Test (BIT) (Wilson, Cockburn & Halligan, 1987) was used as a measure of perception, specifically visuospatial neglect. Visuospatial neglect is a phenomenon commonly associated with stroke in which patients fail to notice, orient, or respond to stimuli in one side of space (the side contralateral to the brain lesion). The Star Cancellation test comprises of a sheet of A4 paper containing 52 large stars, 13 randomly positioned letters and 10 short words, interspersed with 56 smaller stars. The subjects are required to cross out all the small stars. The examiner demonstrates by crossing out 2 stars in the middle section. The maximum score is 54. The normative data presented in Wilson et al. (1987) gives a 52 as a cut-off score for impairment. This is based on an, adequate, sample of 50 non-brain damaged adults, aged 22-82. The star cancellation subtest was chosen to use in isolation from the whole BIT as it has been shown to have excellent diagnostic validity on its own. In a sample of 80 patients, Halligan and Marshall (1989) found that the star cancellation subtest was sensitive to all 30 cases of visual neglect as diagnosed on the basis of the BIT total test score.

Thus, there seemed little benefit to administering the full BIT for the purpose of this study.

1.5.4 Rey-Osterrieth figure copy task

The Rey-Osterrieth Complex Figure copy task (Rey, 1941; Osterrieth, 1944) was used as a second measure of perception. This task requires subjects to copy a complex geometric figure onto a sheet of paper. The specific aspect of perception it measures is visuospatial construction. Visuospatial construction ability is required in activity such as drawing, building and assembling. It requires a combination of visual perception and spatial awareness, as well as more executive skills such as, planning and organisation of motor response. As such, patients with both parieto-occipital lesions and frontal lobe impairment have deficits on the Rey-Osterrieth Complex Figure copy test (Pillon, 1981). (In the current study, visuospatial construction was assessed primarily as an indication of perception, although performance on the task was also considered in the evaluation of executive functioning.) The Rey-Osterrieth Complex Figure copy task has been found to have good concurrent validity, correlating with other visuospatial tests e.g. block design and object assembly from the Wechsler Intelligence tests (Poulton & Moffitt, 1995). Inter-rater reliability has been found to be good (Liberman, Stewart, Seines & Gordon, 1994).

The Rey-Osterrieth Complex Figure test has been presented in a manualised format by Meyers and Meyers (1995), together with a standardised method for scoring and results of a normative standardization study. The norms presented in Meyers and Meyers (1995) were considered suitable for use in the current

study, having been obtained from a large sample which includes a good representation of older adults (n= 601, age range = 18-89 years). The normative scores are stratified by age. (Other demographic variables, gender and years of education, were not found to correlate with performance).

1.5.5 Logical Memory

The Logical Memory subtest from the Wechsler Memory Scales III WMS III) was used to assess verbal memory (Wechsler, 1999). This involves the free recall of information immediately following auditory presentation of a story. In addition to immediate recall, subjects are tested for their retention of the story after a 30 minute delay. Logical Memory has been reported to have good diagnostic sensitivity and is more sensitive than list learning tasks and a paired associate learning tasks at identifying impaired patients from controls (Guilmette & Rasile, 1995). Inter-rater reliability has been reported to be excellent (Woloxzun, Murphy, Wetzel & Fisher, 1993; McGuire & Batchelor, 1998), and test-retest reliability is adequate (The Psychological Corporation, 1997). The best norms available are those presented in the WMS III manual which are based on a sample of 1250 individuals aged 16-89 years (The Psychological Corporation, 1997).

1.5.6 Rey-Osterrieth figure recall

The Rey-Osterrieth figure (Rey, 1941; Osterrieth, 1944) was used to assess non-verbal memory recall, using the administration procedure and norms presented in the manual by Meyers and Meyers (1995). This test requires the subject to draw the figure from memory 3 minutes after the copy procedure

(immediate recall), and again after 30 minutes (delayed recall). A standardised scoring system is used to calculate scores for immediate and delayed recall, based on the number of elements of the figure correctly produced. The Rey-Osterrieth figure has good diagnostic validity, identifying patients with neuropsychological impairment, and loading on a visual memory factor (Berry, Allen, & Schmitt, 1991). Inter-rater, alternate form, test-retest and internal consistency reliability are adequate-good (Berry et al., 1991). Various versions of the complex figure tests exist, but results tend to be comparable. The version presented in Meyers and Meyers (1995) was favoured above others as the corresponding norms are derived from a sample including a good representation of older adults.

1.5.7 Digit Span test

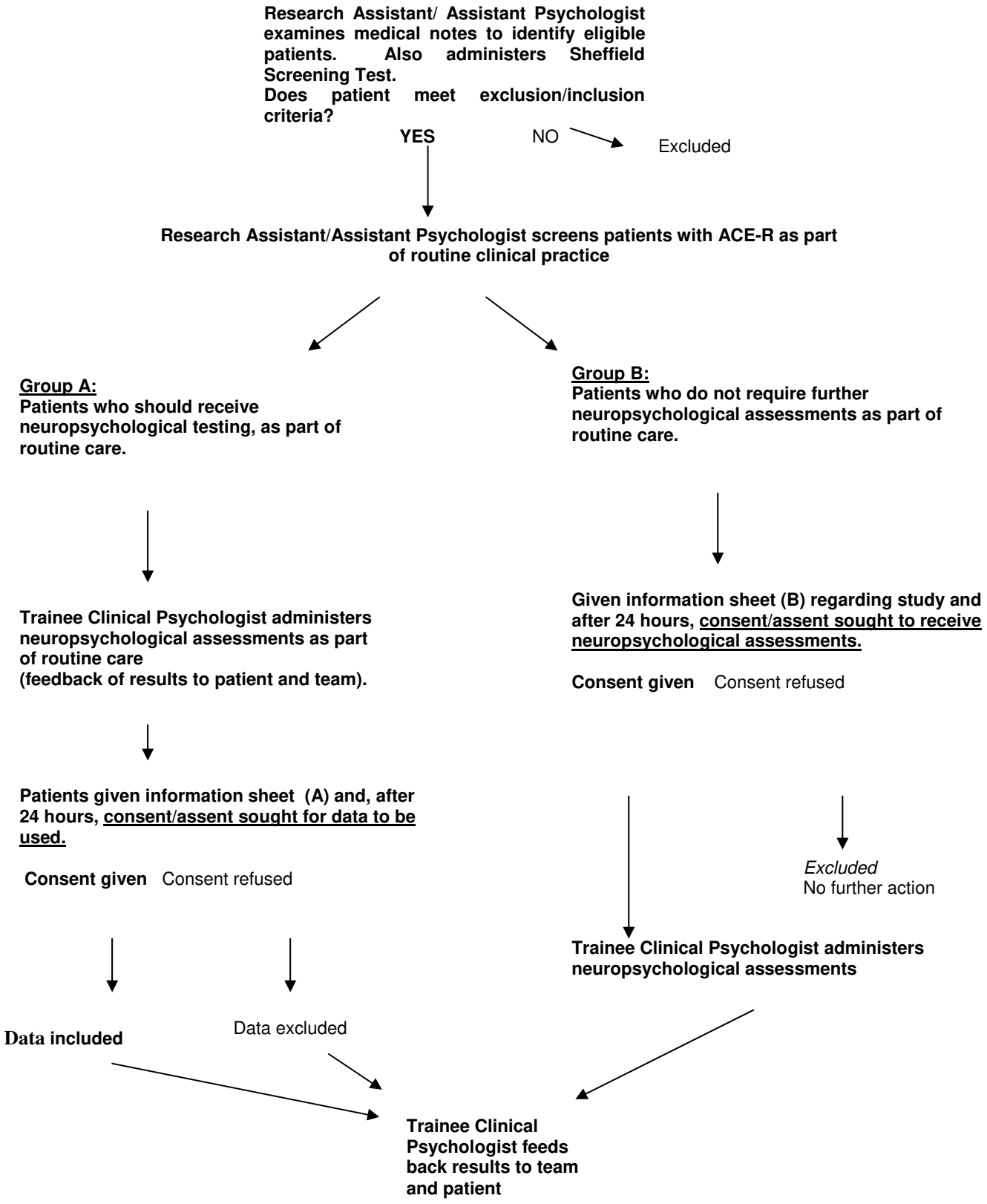
The Digit Span subtest from the WMS-III was used as a measure of attention (Wechsler, 1999). This specifically measures short-term storage capacity, which is one of a number of mechanisms assumed to underlie attentional ability. Digit Span requires subjects to repeat increasing sequences of digits after the examiner, first forwards, then backwards. Digit Span ability is intact in many persons with brain disorders and tends to recover quickly after damage occurs. The test is most sensitive to more severe brain damage, and to left hemisphere damage rather than right hemisphere damage. (Hom & Reitan, 1984).

1.5.8 Letter-Number Sequencing

The Letter-Number Sequencing subtest from the WMS-III was used as a second measure of attention (Wechsler, 1999). This measures another

attentional mechanism, working memory, i.e. the ability to hold information in mind whilst performing a mental operation. In this task, subjects are asked to repeat a series of letters and numbers (increasing in length), giving numbers first in numerical order, followed by letters in alphabetical order. This requires subjects to keep the information in mind long enough to rearrange the order. Studies have shown Letter-Number Sequencing to have adequate criterion validity in brain injury (Donders, Tulsky & Zhu, 2001), and brain imaging studies have demonstrated activation in areas of the brain associated with working memory (Haut, Kuwabara, Leach & Arias, 2000). Scores on Letter-Number Sequencing are strongly correlated with Digit Span scores, suggesting they measure a similar underlying construct (Crowe, 2000).

Appendix 1.6 Study procedure flowchart



Appendix 1.7 Ethical considerations

There were three main ethical considerations.

Potential for emotional distress and fatigue

It was recognised that some patients might experience emotional distress during the administration of neuropsychological assessments if tests led to the patient having insight into previously unrecognised cognitive impairments. The chance of emotional distress was deemed to be low. However, emotional distress was kept to a minimum by making explicit to patients the purpose of carrying out tests and the possibility that they may reveal impairments. Also, by giving patients reassurance that impairments will be explained to them and that the identification of impairments will assist with the provision of appropriate interventions to aide recovery and rehabilitation. In addition to this, there was risk of tests causing patients fatigue as they require concentration and mental effort. Patients were advised that they could discontinue assessments if they were feeling too tired unwell. They were permitted to reconvene tests at a later date.

Confidentiality and storage of data.

Test data and demographic information was made anonymous using a number coding system, and patient identifiable information (e.g. name, address, hospital number) was kept on a separate list to ensure confidentiality. For the duration of the study, the anonymous data was kept in a locked drawer on NHS premises, in the Clinical Psychology office in the Nottingham City Hospital. It

was also stored onto NHS and University computers for analysis. Patient identifiable information was kept in a separate drawer in the Clinical Psychology office. The data will be archived for 7 years after the completion of the study at the University of Lincoln. It will then be destroyed.

Informed Consent

Ethical approval was received for the consent procedure and followed the regulations of the Mental Capacity Act (2005). Patients who were not able to give informed consent, due to cognitive impairment, were not excluded from the study. This is because a study examining the validity of a screening measure needs to include patients with a range of severity of cognitive impairment. For those patients unable to give informed consent, assent was sought from a relative or another person known to the patient. Information sheets were provided to patients (or relatives) with information about the study. These were supplemented with verbal explanations where necessary. The information sheets were written in simple language and large font to assist readability for persons with cognitive impairments or problems with vision. Where necessary, capacity to consent was checked by asking the patient to explain in their own words their understanding of the study and what their participation would entail. Consent or assent was obtained in writing, where possible. If written consent was not possible, e.g. due to physical impairment, verbal consent was witnessed by a member of the psychology or medical staff and this was documented in the medical notes. Consent and assent forms were also filed in the medical file, and patients (or relatives) were given a copy for their reference.

Appendix 2

Additional results

Appendix 2.1 Normality of distributions

Analysis was performed on (continuous) demographic variables and screening test scores to check for normality of distributions.

Demographic variables

Shaprio-Wilk analysis showed that data for age and years of education was not normally distributed; $W(31)= 0.93$, $p<0.05$; $P(32)= 0.70$, $p<0.05$. Data for Barthel scores was normally distributed, $W(32)= 0.98$, $p>0.05$. However, being an ordinal scale, the data from the Barthel Index would also be treated as non-parametric data.

Performance on screening tests

Shapiro-Wilk tests showed that the data for MMSE total scores and ACE-R total scores were both normally distributed; $W(32)= 0.94$, $p>0.05$; $W(32)= 0.94$, $p>0.05$.

The data describing the number of days between stroke and administration of screening tests, and the number of days between stroke and administration of the neuropsychological test battery were not normally distributed; $W(32)= 0.76$, $p<0.05$; $W(36)= 0.79$, $p<0.05$.

Appendix 2.2 Comparison between recruited and non-recruited samples

Analysis was conducted to ensure demographic variables and performance on the screening tests did not differ significantly between patients who were recruited for the full study (neuropsychological battery completed) and those who were not recruited (neuropsychological battery not completed).

Demographic variables

Chi-square tests were used for analysis of categorical variables. These showed that the two samples did not differ significantly in gender, $x^2(1, N= 64)= 0.61$, $p>0.05$, type of stroke, $x^2(3, N=52)= 3.18$, $p>0.05$, hemisphere of stroke, $x^2(1, N=56)= 0.02$, $p>0.05$, presence of hemiparesis, $x^2(1, N=60)= 1.39$, $p>0.05$, or presence of hemianopia, $x^2(1, N=59)= 0.69$, $p>0.05$. Mann-Whitney U tests for nonparametric data were used for analysis of continuous variables. The tests showed that the two samples did not differ significantly in age, U= 416.00, $p>0.05$, Barthel score, U= 286.50, $p>0.05$, or years of education, U= 400.50, $p>0.05$.

Performance on screening tests

Independent samples t-tests showed that there were no significant differences between the two groups in total score on the MMSE, $t(36)= -0.70$, $p>0.05$, or total score on the ACE-R, $t(38)= -0.99$, $p>0.05$. Chi-square analysis showed that there were no significant difference between the two groups in the presence of impairment on the MMSE, as indicated by a cut-off of 24, $x^2(1,$

N=64)= 0.02, *p*>0.05, or at a cut-off of 27, x^2(1, *N*=64)= 1.42, *p*>0.05. There was also no significant difference between the two groups in presence of impairment on the ACE-R, as indicated by a cut-off of 75, x^2(1, *N*=64)= 0.04, *p*>0.05, or a cut-off of 82, x^2(1, *N*=64)= 0.48, *p*>0.05. With presence of impairment indicated by a cut-off of 88, chi-square analysis found a significant difference between the two groups, with more of those in the recruited sample being impaired that those in the non-recruited sample, x^2(1,*N*=64)= 5.25, *p*<0.05. However, this statistic was deemed invalid due to there being no persons with without impairment in the 'recruited' group.

Further to this, a Mann-Whitney U test showed that the two samples did not differ significantly in the number of days between stroke and administration of the screening tests, U= 303.50, *p*>0.05.

Appendix 2.3 Relationships between demographic variables and performance on cognitive assessments

Analysis was conducted to examine the relationships between demographic variables and overall performance on the MMSE, ACE-R, and neuropsychological test battery. Different statistical tests were used for different variables according to the principles of parametric and non-parametric analysis.

Age

Spearman's rank correlation showed that there was no significant relationship between age and total scores on either the MMSE, $r(38)= -0.15$, $p>0.05$, or the ACE-R, $r(38)= -0.26$, $p>0.05$. Mann-Whitney U analysis showed that there was no significant difference in median age, between those patients impaired on the neuropsychological test battery and those who were not impaired, $U= 1.37.50$, $p>0.05$.

Gender

Independent samples t-tests demonstrated that total scores on the screening tests did not differ significantly between male patients and female patients (MMSE: $t(38)= -1.22$, $p>0.05$; ACE-R: $t(38)= -1.22$, $p>0.05$). Chi square analysis showed that gender did not significantly predict whether someone would be impaired on the neuropsychological test battery, $x^2(1, N=40)= 0.30$, $p>0.05$.

Type of stroke

Analysis of variance (ANOVA) demonstrated a significant difference in total scores on the MMSE, according to type of stroke, i.e. LACS, TACS, PACS and POCS, $F(3,30)= 5.5$, $p< 0.05$. There was also a significant difference in total scores on the ACE-R, according to type of stroke, $F(3,30)= 4.16$, $p<0.05$. Chi-square analysis indicated that type of stroke did not significantly predict whether someone would be impaired on the neuropsychological test battery, $x^2(3, N=34)= 6.94$, $p>0.05$. However, the probability for this analysis was approaching significance ($p= 0.07$).

Hemisphere of stroke

Independent samples t-tests demonstrated that total scores on the screening tests did not differ significantly between patients who had left hemisphere stroke and those who had right hemisphere stroke (MMSE: $t(36)= 0.12$, $p>0.05$; ACE-R: $t(36)= -1.12$, $p>0.05$). Chi square analysis showed that hemisphere of stroke did not significantly predict whether someone would be impaired on the neuropsychological test battery, x^2 (1, N=38)= 2.87, $p>0.05$.

Presence of hemiparesis

Independent samples t-tests demonstrated that total scores on the screening tests did not differ significantly between patients who had hemiparesis and those who did not (MMSE: $t(36)=0.30$, $p>0.05$; ACE-R: $t(36)= 0.70$, $p>0.05$). Chi square analysis showed that presence of hemiparesis did not significantly predict whether someone would be impaired on the neuropsychological test battery, $x^2(1, N=38)= 2.99$, $p>0.05$.

Presence of hemianopia

Independent samples t-tests demonstrated that total scores on the screening tests did not differ significantly between patients who had hemianopia and those who did not (MMSE: $t(36)= 1.19$, $p>0.05$; ACE-R: $t(36)= 1.13$, $p>0.05$). Chi square analysis showed that presence of hemianopia did not significantly predict whether someone would be impaired on the neuropsychological test battery, $x^2(1, N=38)= 0.75$, $p>0.05$.

Years of education

Spearman's rank correlation showed that there was no significant relationship between years of education and total scores on either the MMSE, $r(35)= 0.02$, $p>0.05$, or the ACE-R, $r(35)= 0.14$, $p>0.05$. Mann-Whitney U analysis showed that there was no significant difference in median years of education, between those patients impaired on the neuropsychological test battery and those who were not impaired, $U= 97.50$, $p>0.05$.

Score on Barthel Index

Spearman's rank correlation showed that there was no significant relationship between scores on the Barthel Index and total scores on either the MMSE, $r(35)= 0.16$, $p>0.05$, or the ACE-R, $r(35)= 0.11$, $p>0.05$. Mann-Whitney U analysis showed that there was no significant difference in median scores on the Barthel Index, between those patients impaired on the neuropsychological test battery and those who were not impaired, $U= 96.50$, $p>0.05$.

Number of days between stroke and administration of test

Spearman's rank correlation showed that there was no significant relationship between the number of days since stroke and total scores on either the MMSE $r(34)= -0.21$, $p>0.05$, or the ACE-R, $r(34)= -0.20$, $p>0.05$. Mann-Whitney U analysis showed that there was no significant difference in median days since stroke, between those patients impaired on the neuropsychological test battery and those who were not impaired, $U= 99.50$, $p>0.05$.

Summary

Performance on the MMSE, ACE-R and the neuropsychological test battery was not significantly related to age, gender, hemisphere of stroke, presence of hemiparesis, presence of hemianopia, years of education, score on the Barthel Index and the number of days between stroke and administration of test. There was a significant relationship between type of stroke (LACS, TACS, PACS, and POCS) and performance on the MMSE and performance on the ACE-R. The relationship between type of stroke and impairment on the neuropsychological test battery approached significance levels.

Appendix 2.4 Crosstablulations of overall impairment on tests

Tables 1-5 show crosstablulations of number of patients impaired on the screening tests at different cut-offs, and number of patients with overall impairment on the neuropsychological test battery.

Table 1. MMSE (cut-off= 27) and neuropsychological test battery

		Neuropsychological test battery		Total
		Not impaired	Impaired	
MMSE (cut-off= 27)	Not impaired	2	6	8
	Impaired	7	25	32
	Total	9	31	40

Table 2. MMSE (cut-off= 24) and neuropsychological test battery

		Neuropsychological test battery		Total
		Not impaired	Impaired	
MMSE (cut-off= 24)	Not impaired	6	15	21
	Impaired	3	16	19
	Total	9	31	40

Table 3. ACE-R (cut-off= 88) and neuropsychological test battery

		Neuropsychological test battery		Total
		Not impaired	Impaired	
ACE-R (cut-off= 88)	Not impaired	0	0	0
	Impaired	9	31	40
	Total	9	31	40

Table 4. ACE-R (cut-off= 82) and neuropsychological test battery

		Neuropsychological test battery		Total
		Not impaired	Impaired	
ACE-R (cut-off= 82)	Not impaired	4	7	11
	Impaired	5	24	29
	Total	9	31	40

Table 5. ACE-R (cut-off= 75) and neuropsychological test battery

		Neuropsychological test battery		Total
		Not impaired	Impaired	
ACE-R (cut-off= 75)	Not impaired	4	13	17
	Impaired	5	18	23
	Total	9	31	40

Appendix 2.5 MMSE/ACE-R sensitivity and specificity rates

Tables 3 and 4 show MMSE and ACE-R sensitivity and specificity rates at a full range of cut-offs.

Table 3. MMSE sensitivity and specificity rates

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
16.0	0.0	100
17.5	7	100
18.5	7	78
20.0	13	78
21.5	16	78
22.5	29	78
23.5	42	78
24.5	52	67
25.5	61	56
26.5	71	44
27.5	81	22
28.5	90	11
29.5	94	0
31.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Table 4. ACE-R sensitivity and specificity rates

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
37.0	0	100
39.0	3	100
41.5	7	100
47.0	7	99
51.5	10	99
55.0	16	99
60.0	16	78
62.5	23	78
64.0	32	78
65.5	36	78
66.5	42	78
68.0	42	67
69.5	45	67
71.0	48	67
72.5	55	56
74.5	58	44
76.5	61	44
78.0	65	44
79.5	71	44
80.5	74	44
82.0	77	44
83.5	77	33
85.5	81	22
87.5	90	22
89.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Appendix 2.6 Crosstablulations of impairment on ACE-R subscales

Tables 5,6,7 and 8 show crosstablulations of number of patients impaired on the ACE-R subscales screening tests, and number of patients impaired on specific domains on the neuropsychological test battery.

Table 5. ACE-R 'Attention and Orientation' and attention

		Attention		Total
		Not impaired	Impaired	
ACE-R 'Attention and Orientation'	Not impaired	13	0	13
	Impaired	20	4	24
Total		33	4	37

Table 6. ACE-R 'Memory' and memory

		Memory		Total
		Not impaired	Impaired	
ACE-R 'Memory'	Not impaired	5	2	7
	Impaired	11	11	22
Total		16	13	29

Table 7. ACE-R 'Fluency' and executive functioning

		Executive functioning		Total
		Not impaired	Impaired	
ACE-R 'Fluency'	Not impaired	5	4	9
	Impaired	9	18	27
Total		14	22	36

Table 8. ACE-R 'Visuospatial' and executive functioning.

		Perception		Total
		Not impaired	Impaired	
ACE-R 'Visuospatial'	Not impaired	6	3	9
	Impaired	7	20	27
Total		13	23	36

Appendix 2.7 Sensitivity and specificity rates for ACE-R subscales

Tables 9,10,11 and 12 show sensitivity and specificity rates, at a full range of cut-offs, for each ACE-R subscale.

Table 9. ACE-R 'Attention and Orientation'

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
8.0	0	100
9.5	0	97
11.0	25	94
12.5	50	94
13.5	50	88
14.5	75	82
15.5	75	71
16.5	100	47
17.5	100	15
19.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Table 10. ACE-R 'Memory'

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
8.0	00	100
9.5	07	100
11.0	14	88
12.5	21	88
13.5	36	88
14.5	36	56
15.5	43	56
16.5	57	44
17.5	79	44
18.5	86	25
19.5	93	19
21.5	93	13
23.5	100	13
25.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Table 11. ACE-R 'Fluency'

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
2.0	0	100
3.5	27	93
4.5	36	93
5.5	46	87
6.5	59	87
7.5	64	87
8.5	73	53
9.5	86	40
10.5	86	20
11.5	91	7
13.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Table 12. ACE-R 'Visuospatial'

Cut-off less than or equal to:	Sensitivity (%)	Specificity (%)
.5	4	100
2.5	9	100
5.0	13	100
7.0	13	93
9.0	22	86
10.5	35	86
11.5	48	79
12.5	52	71
13.5	70	71
14.5	87	50
15.5	91	14
17.0	100	0

Acceptable levels of sensitivity and specificity are highlighted

Appendix 3

Extended discussion

Appendix 3.1 Suggestions for improved methodology

One of the limitations highlighted for this study was the interval between the administration of the screening tests and the neuropsychological test battery (median= 2 days (IQR 6), which may have resulted in relatively better performance on the neuropsychological test battery, leading to incorrect false positive diagnoses on the screening tests. The reason for this interval was mainly due to practical and organisational factors. Two separate researchers were needed to administer the screening test and the neuropsychological test battery so that the researcher doing the neuropsychological test battery was blind to ACE-R scores. Unfortunately, the two researchers only had a limited amount of time each week for involvement on the project, and were rarely available to administer the tests on the same day. This meant an inevitable delay between the administration of the screen and the battery. On occasions, this delay was enhanced further by a patient being unavailable to complete the neuropsychological battery on the researcher's first visit. If this study were replicated, this problem would ideally be addressed by ensuring the two researchers are available for testing on the same day. Before administering the screening test, the researcher should check with ward staff that the patient is going to be available to complete the neuropsychological test later that day.

A second limitation, identified, was the relatively small proportion of non-impaired patients (nine) compared to impaired patients (31), which may have resulted in poor statistical power and falsely inflated rates of poor specificity. A balanced sample of impaired and non-impaired patients was impossible to achieve through random sampling as base rates of cognitive impairment are so

high in stroke. If the study were replicated, better statistical power could be achieved through a larger sample size. Alternatively patients could be more selectively sampled to ensure a more equal proportion of unimpaired patients. This could be achieved, either through consultation with the ward staff regarding who they feel is least impaired, or by selection of patients according to stroke type (LACS patients are least likely to be impaired). Another possibility is to make use of a group of patients who have had a transient ischemic attack (TIA), which is related in aetiology to a stroke but results in less cognitive impairment.

A third limitation of the study was the high attrition rates between the administration of the screening tests and the neuropsychological test battery, due to patients being discharged home. There is a possibility this may have resulted in a final sample which is poorly representative of acute stroke population with a bias of more impaired patients. If this study were replicated this problem could be addressed by ensuring both the screen and battery are completed on the same day, as mentioned above, or by following up discharged patients at home. Both solutions would require researchers having more time and greater flexibility in their time to give to the project. It would also be preferable if tests were administered during an agreed time frame post-stroke. This would need to be based on clinical evidence regarding the most appropriate time to carry out testing.

Appendix 3.2 Implications for Clinical Psychology practice in stroke services

Both the MMSE and the ACE-R have been found to be invalid as measures for screening for cognitive impairment in acute stroke. There is little justification for their continued use in stroke, especially given that, due to high base rates of cognitive impairment, we can predict cognitive impairment more accurately just by assuming all patients are impaired. Further research may prove successful in identifying a more valid screening measure. However, if, as hypothesised, patients fail screening tests for reasons other than cognitive impairment (i.e. fatigue, physical illness), we would expect to find similar poor validity for all tests. The absence of any valid cognitive screening measures for use in stroke leads to a number of dilemmas regarding how Clinical Psychology should operate in this field.

Presently, pressed Clinical Psychology services rely on screening tests to identify patients who might benefit most from a full neuropsychological assessment. Given it is impossible for Clinical Psychology to administer full neuropsychological assessments to every acute stroke patient, there becomes a need to identify an alternative way of directing these scarce resources. One option would be to identify patients through the *functional* assessments which are routinely carried out on the wards by Occupational Therapists and Physiotherapists, referring those patients who struggle with tasks for reasons other than physical impairment. This could, in fact, prove to be a more effective way of selecting patients because, this way, only patients who are *functionally* affected by cognitive impairment are tested. Thus, the exercise

becomes more practically relevant for patients, and less 'academic'. This would also avoid unnecessarily testing patients whose cognitive impairments are 'trumped' by an inability to carry out ADLs due to serious physical impairments. The problem with this approach is that, following discharge from the ward, patients may continue to recover physically, and cognitive deficits, which were previously unnoticed may emerge. This would also rely on Occupational Therapists, Physiotherapists, and other members of the MDT, being skilled in recognising cognitive influences on practical tasks. Future research could examine the appropriateness of Clinical Psychology referrals based on functional performance, compared to the appropriateness of referrals based on performance on cognitive screening tests.

A further dilemma exists around whether Clinical Psychology services should be assessing patients in acute stroke at all, even with detailed neuropsychological assessments. A number of facts lead us to question whether this practice is justified. Firstly, we have only limited evidence that the more detailed assessments are not invalidated by the same non-cognitive factors as screening measures, as only some assessments have been validated with acute stroke patients. Secondly, even if neuropsychological tests are accurate at describing cognitive impairment, improvement in acute stroke is so rapid that results may be invalidated in a matter of days. Lastly, as referred to above, knowing the extent of cognitive impairment may have little importance in the acute stages post stroke given the extent of physical impairment that many patients experience. There may be more value in waiting to do neuropsychological assessments until after patients have experienced a decent

period of recovery, perhaps even waiting until they have been discharged from hospital. Instead, upon leaving hospital, patients/carers could be given information leaflets regarding cognitive impairment, and be advised to refer for neuropsychological testing if problems persist. However, a possible disadvantage of concentrating resources in the community might be that acute medical services will no longer be able to make use of information about cognitive problems for rehabilitation and discharge planning. It would be useful if further research focused on assessing the value of Clinical Psychology at different stages post-stroke. This should take into account factors such as the relative contribution of assessments to patient well-being, and the predictive validity of assessments regarding recovery and long term difficulties.

Appendix 3.4 Reflections on the study's epistemological position

Quantitative approaches to research are often viewed as being akin to 'positivism', a philosophy of science commonly associated with Auguste Comte (1798-1857). The central tenant of positivism, known as the 'verification principle', is that true statements (e.g. "the ACE-R is a valid measure of cognitive impairment in stroke") are always verifiable by observation and measurement (Smith 1986, cited in Miller, 1999). Thus, scientific research based on this principle, employs quantitative measurement to either prove or disprove hypotheses, ultimately aiming to find a single 'correct' answer to the question.

For the current research it was deemed appropriate to undertake a quantitative (and thus positivistic) approach to the research question. This is because there was a clear need to produce a definitive answer to a very specific clinical question (is the ACE-R a valid instrument for screening for cognitive impairment in stroke?). It was necessary to prove or disprove this statement in order to justify the use of the MMSE and the ACE-R in clinical practice. Whilst a quantitative (positivistic) approach was considered justified, it is important to acknowledge common criticisms of this approach to research.

One key criticism of positivism and quantitative research is directed at the validity of the 'verification principle' itself. This was called into question by Popper (1959, cited in Miller, 1999), who reasoned that it was impossible to ever verify a statement as being correct beyond any possible doubt. This is

because, no matter how many times we find evidence to support a statement, there is always a chance that, in the future, we will find evidence to refute it. Infamously, Popper used the example of the 'black swan' to illustrate this point, saying we can't prove all swans are white, until we have seen a black one. Thus, according to Popper's theory, we cannot be certain that the finding of the current study (the ACE-R is an invalid measure of cognitive impairment in stroke) is true, because it is possible a future study will falsify this finding. Miller (1999), however, suggests that Popper's criticism of pure positivism need not necessarily affect the credibility of quantitative research. Miller (1999) suggests that it is possible to accept the findings of quantitative research, so long as we also acknowledge that this reality may not be wholly accurate and unambiguous. Thus, we can accept that the ACE-R is not a valid measure in stroke, so long as we acknowledge, that this finding is not a certainty, and that the measurement influencing this outcome may have been imperfect.

Another key criticism of positivism and quantitative approaches, especially in respect to Clinical Psychology, is that by investing in the search for a definitive answer to a question, we reduce complex human experience to a set of simplistic indices (Miller, 1999). In this respect, qualitative research is often favoured over quantitative research as outcomes comprise of humanistic descriptions, or sets of descriptions, of the broader phenomenon, rather than a definitive answer to a specific aspect. Miller (1999), however, defends the use of quantitative research in Clinical Psychology, suggesting that reductionism is acceptable so long as we are clear about what specific aspect of a particular phenomenon we are explaining. In relation to the current research, a narrow

focus and definitive (quantitative) answer about the validity of the ACE-R was what was required in order to inform clinical practice. However, it is acknowledged that, given a different clinical focus, qualitative approaches have great value for exploring the experiences of stroke patients in relation to cognitive impairment and their experience of neuropsychological testing.

Appendix 4

Patient information sheets and consent forms

There were two versions of patient information sheets and consent forms. Form A was given to patients who did not need to consent to neuropsychological testing (because completed tests as part of routine care). These patients were only required to give consent for their anonymous data to be used. Form B were given to patients who would not normally be required to undertake detailed neuropsychological testing, and therefore, needed to give consent to complete the tests. Refer to appendix 1.6 for further clarification of the recruitment and consent process.

Nottingham University Hospitals NHS
NHS Trust

Patient Information Sheet (A)

The validity of the Addenbrooke's Cognitive examination- Revised (ACE-R) in Stroke

Investigators : Katie Morris; Dr Vicki Hacker; Prof. Nadina Lincoln

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives or your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

After someone has suffered a stroke, it is important to assess whether they have 'cognitive' problems e.g. problems with memory or language. The aim of this study is to investigate whether a short screening test (ACE-R) is effective in detecting cognitive problems in the early stages after a stroke.

Why have I been chosen?

We are asking 60 people who have had a stroke recently to take part.

What will happen to me if I take part?

You will not have to do anything to take part. You have already had some cognitive assessments, including the ACE-R test, as

part of your routine care. We would just like your permission to include your results in this study.

Do I have to take part?

Taking part is voluntary. It is up to you to decide. If you decide not to take part, your standard of care will not be affected.

What happens to me if I decide not to take part?

Your results will not be included in the study.

What happens if I change my mind?

You can withdraw from the study at any point, without giving a reason. If this happens, the information you have provided will not be included within our study. Your withdrawal will not affect the care you receive.

What are the possible benefits of taking part?

There are no benefits to you. Future stroke patients may benefit as the information may help to improve methods of assessing cognitive problems.

What are the possible disadvantages of taking part?

There are no particular disadvantages of taking part in the study.

What if there is a problem?

You can speak to a member of the research team about any concerns you have. If you wish to make a formal complaint about any aspect of the way you have been treated during the course of this study, you can do this through the NHS Complaints Procedure.

Will my taking part in the study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential and any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised.

What will happen to the results of the study?

They will be presented as a university degree thesis and presented for publication in a scientific journal. Your own scores on the tests will not be personally identifiable in any report. We will send you a summary of results of the study when it is finished (probably in 2009).

Who is organising this study?

The Stroke Service at Nottingham University Hospitals NHS Trust and the University of Lincoln.

Who has reviewed the study?

This study has been reviewed and given favourable opinion by ……..Research Ethics Committee.

What happens now?

Take time to decide whether you want to take part in the study. If you decide to take part, you will be asked to sign a consent form.

Thank you for taking time to read this.

Contact information

If you have any questions about this study you may contact:

Dr Vicki Hacker
Clinical Psychologist
Nottingham University
Hospitals NHS Trust
Beeston Ward
Hucknall Road
Nottingham
NG5 1PB
0115 9691169 ext.46550

Prof. Nadina Lincoln
Research Director
University of Nottingham
William Lee Buildings 8
Science & Technology Park
Nottingham
NG7 2RQ
0115 9515315

Miss Katie Morris
Trainee Clinical Psychologist
University of Lincoln
Health, Life and Social Sciences
Court 11, Satellite Building 8
Brayford Pool
Lincoln
LN6 7TS
01522 886029

Nottingham University Hospitals NHS
NHS Trust

Patient Information Sheet (B)

The validity of the Addenbrooke's Cognitive examination- Revised (ACE-R) in Stroke

Investigators : Katie Morris; Dr Vicki Hacker; Prof. Nadina Lincoln

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives or your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

After someone has suffered a stroke, it is important to assess whether they have 'cognitive' problems e.g. problems with memory or language. The aim of this study is to investigate whether a short screening test (ACE-R) is effective in detecting cognitive problems in the early stages after a stroke.

Why have I been chosen?

We are asking 60 people who have had a stroke recently to take part.

What will happen to me if I take part?

You have already had the ACE-R test. We would like to compare these results with the results of more detailed tests. You will be

asked to complete a number of tasks that assess a range of memory and thinking abilities. These usually take 30-40 minutes to complete. The tests will be completed in a quiet room on the ward with a researcher. If you wish, a carer or another member of staff may be present during the testing. Your results will be explained to yourself and, if you wish, to the medical staff.

Do I have to take part?

Taking part is voluntary. It is up to you to decide. If you decide not to take part, your standard of care will not be affected.

What happens to me if I decide not to take part?

You will not need to complete any more tests at this time. Any further assessments, which you undertake as part of your routine care, will not be included in the study.

What happens if I change my mind?

You can withdraw from the study at any point, without giving a reason. If this happens, the information you have provided will not be included within our study. Your withdrawal will not affect the care you receive.

What are the possible benefits of taking part in the study?

You will receive more detailed cognitive assessments in addition to your routine assessments. This more detailed assessment may allow more accurate detection of cognitive problems. In addition, future stroke patients may benefit as the information may help to improve methods of assessing cognitive problems.

What are the possible disadvantages of taking part?

There are no physical or emotional risks in taking part in this study. It is possible that if you are unable to answer all the questions you may feel as though you are failing. This is not the case as different people perform differently on all aspects of the

test. At the end of the session there will be a chance to discuss your experience and ask any questions.

What if there is a problem?

You can speak to a member of the research team about any concerns you have. If you wish to make a formal complaint about any aspect of the way you have been treated during the course of this study, you can do this through the NHS Complaints Procedure.

Will my taking part in the study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential and any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised.

What will happen to the results of the study?

They will be presented as a university degree thesis and presented for publication in a scientific journal. Your own scores on the tests will not be personally identifiable in any report. We will send you a summary of results of the study when it is finished (probably in 2009).

Who is organising this study?

The Stroke Service at Nottingham University Hospitals NHS Trust and the University of Lincoln.

Who has reviewed the study?

This study has been reviewed and given favourable opinion by ……..Research Ethics Committee.

What happens now?

Take time to decide whether you want to take part in the study. If you decide to take part, you will be asked to sign a consent form.

Thank you for taking time to read this.

Contact information

If you have any questions about this study you may contact:

Dr Vicki Hacker
Clinical Psychologist
Nottingham University
Hospitals NHS Trust
Beeston Ward
Hucknall Road
Nottingham
NG5 1PB
0115 9691169 ext.46550

Prof. Nadina Lincoln
Research Director
University of Nottingham
William Lee Buildings 8
Science & Technology Park
Nottingham
NG7 2RQ
0115 9515315

Miss Katie Morris
Trainee Clinical Psychologist
University of Lincoln
Health, Life and Social Sciences
Court 11, Satellite Building 8
Brayford Pool
Lincoln
LN6 7TS
01522 886029

Nottingham University Hospitals NHS
NHS Trust

Patient Number:

CONSENT FORM (A)

The validity of the Addenbrooke's Cognitive examination- Revised (ACE-R) in Stroke

Investigators: Dr Vicki Hacker, Prof. Nadina Lincoln, Miss Katie Morris

1) I have read and understood the information sheet (dated 06/08/07, version 1). ☐

2) My participation is voluntary. I can withdraw at any time, without giving any reason. ☐

3) I agree to take part in the above study. ☐

4) The researchers can have access to my medical records. ☐

Name of Patient	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature
_____	_____	_____

1 copy for patient, 1 for research file; 1 (original) to be kept in medical notes).

Patient Number:

CONSENT FORM (B)

The validity of the Addenbrooke's Cognitive examination- Revised (ACE-R) in Stroke

Investigators: Dr Vicki Hacker, Prof. Nadina Lincoln, Miss Katie Morris

5) I have read and understood the information sheet (dated 06/08/07, version 1). ☐

6) My participation is voluntary. I can withdraw at any time, without giving any reason. ☐

7) I agree to take part in the above study. ☐

8) The researchers can have access to my medical records. ☐

5) I <u>would/ would not</u> like the results of my assessments to be passed to the medical team (please delete as appropriate).

Name of Patient	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature
_____	_____	_____

1 copy for patient, 1 for research file; 1 (original) to be kept in medical notes).

Appendix 5

The ACE-R

(Mioshi et al., 2006)

ADDENBROOKE'S COGNITIVE EXAMINATION - ACE-R

Final Revised Version A (2005)

Name :
Date of birth :
Hospital no. :

Addressograph

Date of testing: ..... /..... /.....
Tester's name: .....
Age at leaving full-time education: .....
Occupation: .....
Handedness: .....

ORIENTATION

➢ Ask: What is the	Day	Date	Month	Year	Season	[Score 0-5]
➢ Ask: Which	Building	Floor	Town	County	Country	[Score 0-5]

REGISTRATION

➢ Tell: 'I'm going to give you three words and i'd like you to repeat after me: lemon, key and ball'. After subject repeats, say 'Try to remember them because i'm going to ask you later'. Score only the first trial (repeat 3 times if necessary).

Register number of trials .....

[Score 0-3]

ATTENTION & CONCENTRATION

➢ Ask the subject: ' could you take 7 away from a 100? After the subject responds, ask him or her to take away another 7 to a total of 5 subtractions. If subject make a mistake, carry on and check the subsequent answer (i.e. 93, 84, 77, 70, 63 -score 4)

Stop after five subtractions (93, 86, 79, 72, 65). ....

[Score 0-5]
(for the best performed task)

➢ Ask: 'could you please spell **WORLD** for me? Then ask him/her to spell it backwards:

....

MEMORY - Recall

➢ Ask: 'Which 3 words did I ask you to repeat and remember?'

....

[Score 0-3]

MEMORY - Anterograde Memory

➢ Tell: ' I'm going to give you a name and address and I'd like you to repeat after me. We'll be doing that 3 times, so you have a chance to learn it. I'll be asking you later'

Score only the third trial

[Score 0-7]

	1st Trial	2nd Trial	3rd Trial
Harry Barnes	....	....	....
73 Orchard Close	....	....	....
Kingsbridge	.....	.....	.....
Devon	.....	.....	.....

MEMORY - Retrograde Memory

[Score 0 -4]

➢ Name of current Prime Minister .....
➢ Name of the woman who was Prime Minister .....
➢ Name of the USA president .....
➢ Name of the USA president who was assassinated in the 1960's .....

ATTENTION&ORIENTATION MEMORY

VERBAL FLUENCY - Letter 'P' and animals

➢ **Letters**

Say: 'I'm going to give you a letter of the alphabet and I'd like you to generate as many words as you can beginning with that letter, but not names of people or places. Are you ready? You've got a minute and the letter is P'

[Score 0 - 7] ☐

>17	7
14-17	6
11-13	5
8-10	4
6-7	3
4-5	2
2-3	1
<2	0
total	correct

➢ **Animals**

Say: 'Now can you name as many animals as possible, beginning with any letter?

[Score 0 - 7] ☐

>21	7
17-21	6
14-16	5
11-13	4
9-10	3
7-8	2
5-6	1
<5	0
total	correct

FLUENCY

LANGUAGE - Comprehension

➢ Show written instruction:

[Score 0-1] ☐☐

Close your eyes

➢ 3 stage command:
'Take the paper in your right hand. Fold the paper in half. Put the paper on the floor'

[Score 0-3] ☐☐

LANGUAGE - Writing

➢ Ask the subject to make up a sentence and write it in the space below:
Score 1 if sentence contains a subject and a verb (see guide for examples)

[Score 0-1] ☐☐

LANGUAGE

LANGUAGE - Repetition

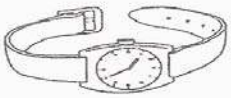
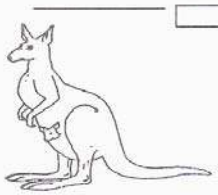
➢ Ask the subject to repeat:' **hippopotamus'; 'eccentricity; 'unintelligible'; 'statistician'**
Score 2 if all correct; 1 if 3 correct; 0 if 2 or less. [Score 0-2] ☐

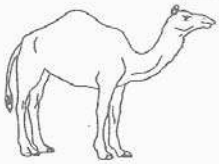
➢ Ask the subject to repeat: **'Above, beyond and below'** [Score 0-1] ☐

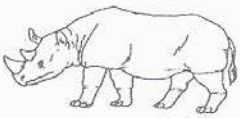
➢ Ask the subject to repeat: **'No ifs, ands or buts'** [Score 0-1] ☐☐

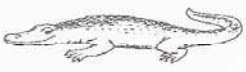
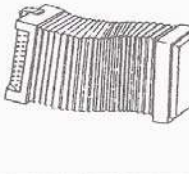
LANGUAGE - Naming

➢ Ask the subject to name the following pictures: [Score 0-2] pencil + watch ☐☐



[Score 0-10] ☐

LANGUAGE - Comprehension

➢ Using the pictures above, ask the subject to: [Score 0-4] ☐

- Point to the one which is associated with the monarchy _____
- Point to the one which is a marsupial _____
- Point to the one which is found in the Antarctic _____
- Point to the one which has a nautical connection _____

LANGUAGE

ADDENBROOKE'S COGNITIVE EXAMINATION - ACE-R Final Revised Version (2005)

LANGUAGE - Reading

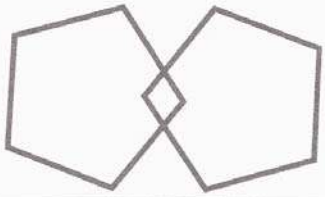
➢ Ask the subject to read the following words: [Score 1 only if all correct] [Score 0-1] ☐

sew
pint
soot
dough
height

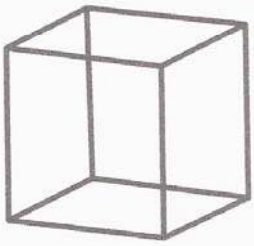
LANGUAGE

VISUOSPATIAL ABILITIES

➢ Overlapping pentagons: Ask the subject to copy this diagram: [Score 0-1] ☐☐



➢ Wire cube : Ask the subject to copy this drawing (for scoring, see instructions guide) [Score 0-2] ☐



➢ Clock: Ask the subject to draw a clock face with numbers and the hands at ten past five. (for scoring see instruction guide: circle = 1, numbers = 2, hands = 2 if all correct) [Score 0-5] ☐

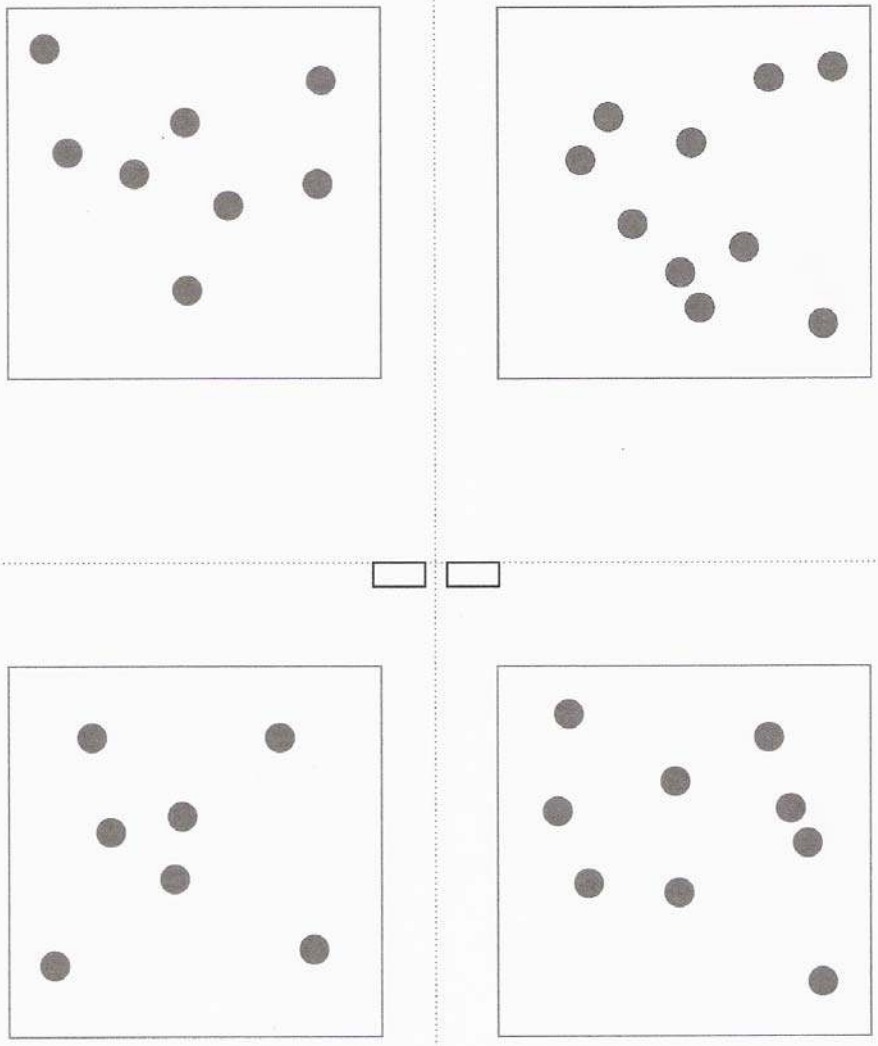
VISUOSPATIAL

PERCEPTUAL ABILITIES

VISUOSPATIAL

- Ask the subject to count the dots without pointing them

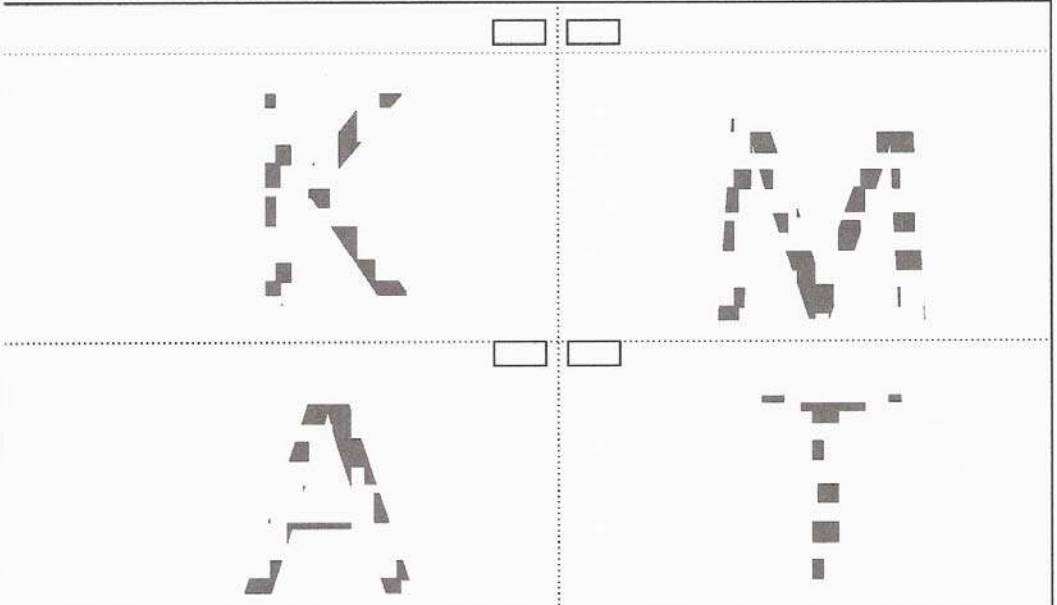
[Score 0-4]



ADDENBROOKE'S COGNITIVE EXAMINATION - ACE-R

PERCEPTUAL ABILITIES

➢ Ask the subject to identify the letters [Score 0-4]



RECALL

➢ Ask "Now tell me what you remember of that name and address we were repeating at the beginning"

[Score 0-7]

Harry	Barnes	....
73	Orchard	Close ....
	Kingsbridge	.....
	Devon	.....

RECOGNITION

➢ This test should be done if subject failed to recall one or more items. If all items were recalled, skip the test and score 5. If only part is recalled start by ticking items recalled in the shadowed column on the right hand side. Then test not recalled items by telling "ok, I'll give you some hints: was the name X, Y or Z?" and so on. Each recognised item scores one point which is added to the point gained by recalling.

[Score 0-5]

Jerry Barnes	Harry Barnes	Harry Bradford	recalled
37	73	76	recalled
Orchard Place	Oak Close	Orchard Close	recalled
Oakhampton	Kingsbridge	Dartington	recalled
Devon	Dorset	Somerset	recalled

General Scores

MMSE	/30
ACE-R	/100

Subscores

Attention and Orientation	/18
Memory	/26
Fluency	/14
Language	/26
Visuospatial	/16

VISUOSPATIAL MEMORY SCORE

Normative values based on 63 controls aged 52-75 and 142 dementia patients aged 46-86

Cut-off <88 gives 94% senstivity and 89% specificity for dementia
Cut-off <82 gives 84% sensitivity and 100% specificity for dementia

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual for Mental and Behavioural Disorders, 4th edition*, Washington DC: American Psychiatric Association.

Appelros, P. & Andersson, A.G. (2006). Changes in Mini Mental State Examination score after stroke: Lacunar infarction predicts cognitive decline. *European Journal of Neurology, 13*(5): 491-495.

Berg, E. (1948). *A* simple objective treatment for measuring flexibility in thinking. *Journal of General Psychology, 39*, 15-22.

Berry, D.T., Allen, R.S., & Schmitt, F.A. (1991). Rey-Osterrieth Complex Figure: Psychometric characteristics in a geriatric sample. *Clinical Neuropsychologist, 5*(2), 143-153.

Bielak, A.M., Mansueti, L., Strauss, E., & Dixon, R.A. (2006). Performance on the Hayling and Brixton tests in older adults: Norms and correlates. *Archives of Clinical Neuropsychology, 21*(2), 141-149.

Bieliauskas, L.A., Depp, C., Kauszler, M.L., Steinberg, B.A., & Lacy, M. (2000). IQ scores on the Mini-Mental State Examination (MMSE). *Aging, Neuropsychology and Cognition, 7*(4), 227-229.

Blake, H., McKinney, M., Treece, K., Lee, E., & Lincoln, N. (2002). An evaluation of screening measure for cognitive impairment after stroke. *Age and Ageing, 31*(6*)*, 451-456.

Blessed, G., Black, S.E., Butler, T., & Kay, D.W. (1991). The diagnosis of dementia in the elderly. A comparison of CAMCOG (the cognitive section of CAMDEX), the AGECAT progra, DSM-III, the Mini-Mental State Examination and some short rating scales. *The British Journal of Psychiatry: the Journal of Mental Science, 159*, 193-198.

Burgess, P.W. & Shallice, T. (1997). *Hayling and Brixton Tests*. Oxford: Harcourt Assessment.

Claesson, L., Linden, T., Skoog, I., & Blomstrand, C. (2005). Cognitive impairment after stroke- impact on activities of daily living and costs of care for elderly people. *Cerebrovascular Diseases, 19*(2), 102-109.

Crowe, S.F. (2000). Does the letter number sequence task measure anything more than digit span? *Assessment, 7*(2), 113-117.

Darvesh, S., & Freedman, M. (1996). Subcortical dementia: a neurobehavioral approach. *Brain and Cognition, 31*, 230–249.

De-Koning, I., van-Kooten, F., Dippel, D., van-Harskamp, F., Grobbee, D., Kluft, C., & Koudstaal, P. (1998). The CAMCOG: a useful screening instrument for dementia in stroke patients. *Stroke, 29*(10), 2080-2086.

De-Koning, I., Dippel, D., van-Kooten, F., & Koudastall, P. (2000). A short screening instrument for post stroke dementia: the R-CAMCOG. *Stroke*, *31*(7), 1502-1508.

De-Koning, I., van-Kooten, F., Koudstaal, P.J., & Dippel, D.W.J. (2005). Diagnostic value of the Rotterdam-CAMCOG in post-stroke dementia. *Journal of Neurology, Neurosurgery and Psychiatry, 76*(2), 263-265.

Dick, J. P., Guiloff, R. J., Stewart, A., Blackstock, J., Bielawska, C., Paul, E. A., et al. (1994). Mini-mental state examination in neurological patients. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 496-499.

Donders, J., Tulsky, D.S., & Zhu, J. (2001). Criterion validity of new WAIS-III subtest scores after traumatic brain injury. *Journal of the International Neuropsychological Society, 7*(7), 892-898.

Dudas, R.B. (2005). The Addenbrooke's Cognitive Examination (ACE) in the Differential Diagnosis of Early Dementias versus affective disorder. *The American Journal of Geriatric Psychiatry, 13*(3), 218-226.

Engstad, T., Viitanen, M., & Almkvist, O. (2007). Cognitive impairment after stroke- diagnosis and management. *Tidsskrift for den Norske Legeforening, 127*(10), 1390-1393.

Folstein, M., Folstein, S., & McHugh, P. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 78-85.

Gaber, T.A.Z.K. (2008). Evaluation of the Addenbrooke's Cognitive Examination's validity in a brain injury rehabilitation setting. *Brain Injury, 22*(7),589-593.

Giordani, B., Boivin, M.J., Hall, A.L., & Foster, N.L. (1990). The utility and generality of Mini-Mental State Examination scores in Alzheimer’s disease. *Neurology*, *40*(12), 1894-1896.

Grace, J., Nadler, J. D., White, D. A., Guilmette, T. J., Giuliano, A. J., Monsch, A. U., et al. (1995). Folstein vs. modified Mini-Mental State Examination in geriatric stroke. Stability, validity, and screening utility. *Archives of Neurology*, *52*, 477-484.

Guilmette, T.J., & Rasile, D. (1995). Sensitivity, specificity and diagnostic accuracy of three verbal memory measures in the assessment of mild brain injury. *Neuropsychology, 9*(3), 338-344.

Halligan, P.W., & Marshall, J.C. (1989). Visuospatial neglect: underlying factors and test sensitivity. *Lancet, 2*, 908-911.

Haut, M.W., Kuwabara, H., Leach, S., & Arias, R.G. (2001). Neural activation during performance of number-letter sequencing. *Applied Neuropsychology, 7*(4), 237-242.

Helkala, E.L., Kivipelto, M., Hallikainene, M., Alhainen, K., Heinonen, H., Tuomilehto, J., Soininen, H., & Nissinen, A. (2002). Usefulness of repeated presentation of Mini-Mental State Examination as a diagnostic procedure: A population-based study. *Acta Neurologica Scandinavica*, *106*(6), 341-346.

Henry, J.D., & Crawford, J.R. (2004). A Meta-Analytic Review of Verbal Fluency Performance Following Focal Cortical Lesions. *Neuropsychology, 18*(2), 284-295.

Hom, J., & Reitan, R.M. (1984). Neuropsychological correlates of rapidly vs. slowly growing intrinsic cerebral neoplasms. *Journal of Clinical and Experimental Neuropsychology, 6*, 309-324.

Huppert, F., Brayne, C., Gill, C., Paykel, E., & Beardsall, L. (1995). CAMCOG- a concise neuropsychological test to assist dementia diagnosis: Socio-demographic determinants in an elderly population sample. *British Journal of Clinical Psychology*, 34, 529-541.

Larner, A.J. (2005). An audit of the Addenbrooke's Cognitive Examination (ACE) in clinical practice. *International Journal of Geriatric Psychiatry, 20*, 593-594.

Larner, A.J. (2006). Audit of the Addenbrooke's Cognitive Examination (ACE) in clinical practice 2: Longitudinal change. *International Journal of Geriatric Psychiatry, 21*, 698-701.

Larner, A.J. (2007). Addenbrooke's Cognitive Examination-Revised(ACE-R) in day-to-day clinical practice. *Age Ageing, 36,* 685–6.

Lesniak, M., Bak, T., Czepiel, W., Seniow, J., & Czlonkowska, A. (2008). Frequency and prognostic value of cognitive disorders in stroke patients. *Dementia and geriatric cognitive disorders, 26*(4), 356-363.

Lezak, M.D., Howieson, DB., & Loring D.W. (2005). *Neuropsychological Assessment (4th ed.).* Oxford: Oxford University Press.

Liberman, J.N., Stewart, W., Seines, O., & Gordon, B. (1994). Rater agreement for the Rey-Osterrieth Complex Figure Test. *Journal of Clinical Psychology, 50*(4), 615-624.

Lincoln, N.B., Nicholl, C.R., & Flannaghan, T. (2003). The validity of questionnaire measures for assessing depression after stroke. *Clinical Rehabilitation, 17*, 840-846.

Loong, T.W. (2003). Understanding sensitivity and specificity with the right side of the brain. *British Medical Journal, 327,* 716-719.

Madureira, S., Guerreiro, M., & Ferro, J.M. (2001). Dementia and cognitive impairment three months after stroke. *European Journal of Neurology, 8* (6), 621-627.

Mathuranath, P., Nestor, P., Berrios, G., Rakowicz, W., & Hodges, J. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology, 55*(11), 1613-1620.

McGuire, B.E., & Batchelor, J. (1998). Inter-rater reliability of the WMS-R Logical Memory and Visual Reproduction subtests in a neurosurgical sample. *Australian Psychologist, 33*(3), 231-233.

Mental Capacity Act of 2005, No.9 (2005). Retrieved November 20, 2008, from http://www.opsi.gov.uk/acts/acts2005/pdf/ukpga_20050009_en.pdf

Meyers, J., & Meyers, K. (1995). *Rey Complex Figure Test and Recognition Trial (RCFT)*. Oxford: Harcourt Assessment.

Miller, E. (1999). Positivism and Clinical Psychology. *Clinical Psychology and Psychotherapy, 6*, 1-6.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21*(11), 1078-1085.

Nadler, J.D., Richardson, E.D., & Malloy, P.F. (1994). Detection of impairment with the Mini-Mental State Examination. *Neuropsychiatry, Neuropsychology, & Behavioural Neurology, 7*(2), 109-113.

Naugle, R.I., & Kawczak K. (1989). Limitations of the Mini-Mental State Examination. *Cleveland Clinic Journal of Medicine, 56,* 277–281.

Nys, G., van-Zandvoort, M., de-Kort, P., Jansen, B., Kappelle, L. & De-Haan, E. (2005a). Restrictions of the Mini-Mental State Examination in acute stroke. *Archives of Clinical Neuropsychology, 20*(5), 623-629.

Nys, G.M.S., van-Zandvoort, M.J.E., de-Kort, P.L.M., Jansen, B.P.W., van-der-Worp, H.B., Kappelle, L.J., & de-Haan, E.H.F. (2005b). Domain-specific cognitive recovery after first-ever stroke: A follow-up study of 111 cases. *Journal of the International Neuropsychological Society, 11* (7), 795-806.

Nys, G.M.S., van-Zandvoort, M.J.E., de-Kort, P.L.M., van-der-Worp, H.B., Jansen, B.P.W., Algra, A., de-Haan, E..H.F., & Kappelle, L.J. (2005c). The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. *Neurology, 64*(5), 821-827.

Nys, G.M.S., van-Zandwoort, M.J.E., de-Kort, P.L.M., Jansen, B.P.W., de-Haan, E.H.F., & Kapelle, L.J. (2007). Cognitive disorders in acute stroke: prevalence and clinical determinants. *Cerebrovascular Diseases, 23*(5-6), 408-416.

Osterrieth, P.A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30*, 206-356.

Patel, M., Coshall, C., Rudd, A.G., & Wolfe, C.D.A. (2003). Natural history of cognitive impairment after stroke and factors associated with its recovery. *Clinical Rehabilitation, 17*(2), 158-166.

Pillon, B. (1981). Negligence de l'hemi-espace gauche dans des epreuves visuo-constructives. *Neuropsychologia, 19*, 317-320.

Poulton, R.G. & Mottitt, T.E. (1995). The Rey-Osterrieth Complex Figure test: norms for young adolescents and an examination of validity. *Archives of Clinical Neuropsychology, 10*(1), 47-56.

Psychological Corporation (The) (1997). *WAIS-III and WMS-III technical manual.* San Antonio, TX: The Psychological Corporation.

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie, 28*, 286-340.

Roth, M., Tym, E., Moutiov, C.Q., Huppert, F.A., Hendrie, H, Verma, S., & Goddard, R. (1986). CAMDEX: A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *The British Journal of Psychiatry: the Journal of Mental Science, 149*, 698-709.

Royal College of Physicians of London. (2004). *National Clinical Guidelines for Stroke* (2nd ed.) [brochure]. London: Authors.

Shulman, K.I., Hermann, N., Brodaty, H., et al. (2006). IPA survey of brief cognitive screening instruments. *International Psychogeriatrics, 18*, 281–94.

Starr, J.M., & Lonie, J. (2007). The influence of pre-morbid IQ on Mini Mental State Examination score at time of dementia presentation. *International Journal of Geriatric Psychiatry, 22*(4), 382-384.

Stroke Association (The). (2006). *What is a stroke?* [fact sheet]. Hampshire (UK): Scriptographic Publications Ltd.

Syder D, Body R, Parker M & Boddy M (1993). *Sheffield Screening Test for Acquired Language Disorders*. Windsor: Nfer-Nelson.

Tierney, M.C., Szalai, J.P., Snow, W.G., Fisher, F.H., & Dunn, E. Domain specificity of the subtests of the Mini-Mental State Examination. *Archives of Neurology, 54*(6), 713-716.

Tombaugh, T.N., & McIntyre, N.J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society, 40*(9), 922-935.

Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*, 167-177.

Trenkle, D.L., Shankle, W.R., & Azen, S.P. (2007). Detecting cognitive impairment in primary care: performance assessment of three screening instruments. *Journal of Alzheimer's Disease, 11*(3), 323-335.

Wechsler, D. (1999). *Wechsler Memory Scale, third edition (WMS-III)*. Oxford: Harcourt Assessment.

Wilson, B., Cockburn, J., & Halligan, P. (1987). Behavioural Inattention Test (BIT). Oxford: Harcourt Assessment.

Woloszyn, D.B., Murphy, S.G., Wetzel, L., & Fisher, W. (1993). Inter-rater agreement on the Wechsler Memory Scale-Revised in a mixed clinical population. *Clinical Neuropsychologist, 7*(4), 467-471.

World Heath Organisation. (1978). *Cerebrovascular disorder: A clinical & research classification, 43.* Geneva: World Health Organisation.

Yang, C., Hwang, J.P., Jen, T.S., & Liu, C.M. (2000). The clinical applications of Mini-Mental State Examination in geropsychiatric inpatients. *International Journal of Psychiatry in Medicine, 30(*3), 227-285.

Word count

Sections 1-3 (paper) = 5, 388

Sections 4-6 = 18,212

Total word count = 23,600