UNIVERSITY OF NOTTINGHAM
INSTITUTE OF WORK, HEALTH AND ORGANISATIONS

The assessment of dementia severity using non-verbal cognitive tests

Sobia Tbsum Khan BSc (Hons)

Submitted to the University of Nottingham in part fulfilment of the requirements for the degree of

Doctor of Clinical Psychology

September 2010

Table of contents

Contents

Abstract

Background

In 2006 the National Institute for Health and Clinical Excellence (NICE) made a controversial decision to restrict the availability of Anti-cholinesterase Inhibitors (AchI) to patients with Alzheimer's Disease (AD) who score within the 'moderate' range (10-20 points) on the Mini-Mental State Examination (MMSE). A high court judge found NICE to have acted illegally by not providing specific guidance for individuals whose first language was not English, hence for whom the MMSE was not administrable. There is currently a lack of alternative objective measures of cognitive functioning that are equivalent to the MMSE, which can be used for people who do not speak English. This could result in inequalities within assessment and treatment practices.

Aim

The aim of this study was to explore whether cognitive tests that did not require spoken English (by patients) could provide equivalent information to that obtained from the MMSE. The second aim was to explore if these alternative cognitive tests could differentiate between those who were eligible for treatment with AchI from those who were not. This research would provide preliminary data about the utility of the chosen tests to decide on treatment eligibility based on the cut-offs on the MMSE proposed by NICE. This would inform research in the future which would select non-English speaking samples to explore the cultural equivalence of the measures selected for this study.

Objectives

The two objectives of this study were firstly to explore the correlation between participants' total scores on the MMSE and their total scores on the Rey Complex Figure Test (RCFT), Raven's Colour Progressive Matrices (RCPM), Symbol Digit Modalities Test (SDMT), Brixton test,

Clock Drawing Test (CDT), and the Colour Trails Test (CTT). Secondly to explore if participants' performance on each of the alternative cognitive tests, namely the RCFT, RCPM, SDMT, Brixton test, CDT or the CTT would differentiate those eligible for treatment with AchI (10-20 score on MMSE) from those who were not.

Methods

Twenty participants (aged 65-90 years), whose first language was English, were recruited from two older people's Mental Health Service sites based in two cities in England. All participants were assessed as having the capacity to consent using a structured assessment of capacity. The seven cognitive tests were administered to all participants. The duration of the testing ranged from 1 hour to an hour and half.

Results

The relationship between the MMSE and the six alternative cognitive tests was assessed using correlation analysis. There was a statistically significant linear relationship between the MMSE and the RCFT visual construction trial (r= .609; P<.006), the RCFT recognition trial (r= .496; P<.031), RCPM (r= .452; p<.045), the SDMT (r=.670; P<.001), the CDT (r=-.577; P<.008) and the CTT 1 (rho=-.576; P<.012).

In order to assess whether or not measures were able to identify those eligible for treatment or not, the measures that significantly correlated with the MMSE, were further analysed using Receiver Operating Characteristic (ROC) analysis. The area under the ROC curve values were as follows: RCFT visual construction (0.750, 95%, CI .524 - .976), RCFT recognition memory (0.801, 95%, CI .590 – 1.012), RCPM (0.573, 95%, CI 0.298 – 0.848), SDMT (0.708, 95%, CI 0.469 – 0.947), CTT1 (0.818, 95%, CI 0.610 – 1.027) and the CDT (0.734, 95%, CI 0.479 – 0.990). All AUC values indicated that

the above measures had moderate to high accuracy apart from the RCPM, which had accuracy that was equal to chance.

Cut-off scores with adequate sensitivity and specificity were identified for all the above measure apart from the RCPM. The cut-off scores with their respective sensitivity and specificity were: RCFT Visual constructions <20.5 (sen 87%, spec 64%); RCFT recognition memory <14.5 (sen 87%, 72%); SDMT <11 (sen 75%, spec 66%); CTT1 > 144 seconds (sen 86%, spec 64%) and the CDT >9 (sen 75%, spec 75%).

Conclusion

Cognitive tests that do not require spoken language had adequate predictive value and have utility in identifying those who are and are not eligible for treatment with AchI. The unequal prevalence rates of positive cases in this sample resulted in reduced PPV values for all the measures. An additional finding was that 80% of the participants in this sample who were not eligible for treatment with AchI according to NICE guidelines, were being treated with an AchI. Further research into the cross-cultural equivalence of the selected tests is necessary.

Statement of contribution

The lead researcher was responsible for the literature review, study design, recruitment, data collection, scoring of tests, planning and conducting the analysis and writing the thesis. Dr Roshan Das Nair provided supervision of the research, academic work and analysis, read drafts and provided feedback for the thesis and the journal paper.

Dr David Connelly and Dr Helen Philpott provided clinical supervision, reviewed the study design and methodology, and provided feedback on written work. They also networked with key collaborators within the area to generate interest and facilitate recruitment.

Professor Nadina Lincoln provided guidance on the application of Receiver Operating Characteristic analysis.

Acknowledgments

First and foremost, I would like to thank the participants who took part in this research. Their enthusiasm and patience during testing was greatly appreciated. Aside from helping to collect data they were a constant reminder of the importance of research in this field, with a view to developing treatment practices for people with dementia.

I would like to thank the following people for their helpful advice: Dr Sally Damms, Dr Barry Greatorex, Dr Chester, Dr Parker and Corinne Gale. I would also like to acknowledge the professionals who referred participants and invited me to present to their teams, in particular Paul Carruthers, Debbie Carter, Pat Sumner, Dr Ola Junaid, Dr Khinde Junaid, Dr Anthony Prakash, Wendy Stevenson and Susan Stayner.

Finally, Baji Adeela, Pops, Shaks, Fozi, Anissa, Baji Asrah and Paijaan, thank you for your patience when I have been unavailable and for all you have done to help me function outside of work. I will make it up to you. Mum and Dad, thank you for everything, this doctorate is a credit to you both.

Journal article

Note: This paper was written for the International Journal of Geriatric Psychiatry. Please see extended paper Appendix 1 for an overview of the criteria for submission. This journal was selected over alternative journals because it has an impact factor of 2.128. In addition, it is specific to research with older people and had an international readership. Therefore, this may increase the likelihood of attracting readers who are interested in the cross-cultural assessment of dementia.

Cognitive assessment of Dementia severity using non-verbal cognitive tests

Sobia Tbsum Khan [a], Roshan Das Nair [a], David Connelly [b], Helen Philpott [c], Nadina Lincoln [a].

[a] University of Nottingham UK; [b] Nottinghamshire NHS Trust UK; [c] Derbyshire Mental Health Services Trust UK

ABSTRACT

Objective: To explore the utility of cognitive tests, that did not rely on spoken language from participants for decision making about eligibility for treatment with Anti-cholinesterase Inhibitors (AchI).

Methods: A cross-sectional design was used. Data was collected from 20 participants, aged 65 to 90 years (M, 77.6: SD, 7.2), with dementia, or cognitive difficulties. The sample comprised of eleven males and nine females. The Mini-Mental State Examination (MMSE), and six alternative cognitive tests were administered, these were; the Rey Complex Figure Test (RCFT), Ravens Colour Progressive Matrices (RCPM), Symbol Digit Modalities Test (SDMT), Brixton Test, Clock Drawing Test (CDT), and Colour Trails Test (CTT).

Results:

There was statistically significant correlations between the MMSE and the following cognitive tests: RCFT visual construction subtest (r= .609; P<.006), the RCFT recognition subtest (r= .496; P<.031),

RCPM (r= .452; p<.045), the SDMT (r=.670; P<.001), the CTT 1 (Rho=-.576; P<.012) and the CDT (r=-.577; P<.008).

The area under the ROC curve values were as follows: RCFT visual construction (0.750, 95%, CI .524 - .976), RCFT recognition memory (0.801, 95%, CI .590 – 1.012), RCPM (0.573, 95%, CI 0.298 – 0.848), SDMT (0.708, 95%, CI 0.469 – 0.947), CTT1 (0.818, 95%, CI 0.610 – 1.027) and the CDT (0.734, 95%, CI 0.479 – 0.990). Cut-off scores with adequate sensitivity and specificity were identified for all the above measures apart from the RCPM, which had predictive accuracy that was equal to chance.

Conclusions

Cognitive tests which do not require spoken language have utility in differentiating between those who are and are not eligible for treatment with AchI, as defined by the cut-off (<20) on the MMSE in the National Institute for Clinical Excellence guidelines (NICE, 2007).

Key words: Anti-cholinesterase Inhibitors, Dementia, NICE, MMSE, 'Non-verbal cognitive assessment, non-English speakers.

Details for correspondence to lead author: Sobia Khan, Trent Doctorate in Clinical Psychology, Institute of Work Health and Organisations, University of Nottingham, International House, B Floor, Jubilee Campus, Nottingham NG8 1BB. Email: lwxstk@nottingham.ac.uk.

INTRODUCTION

Dementia is an acquired degenerative neurological condition, which manifests itself in progressive deterioration of global cognitive functioning, activities of daily living and behaviour; with significant consequences for families and carers (Green, 2000; Rockwood *et al.,* 2007). For the most part, their aetiology is unknown or only partially understood (Lezak *et al.,* 2004; Robillard, 2007). Even with recommended 'gold standard' diagnostic criteria, cases are under-diagnosed, missed or misclassified (Cairns *et al.,* 2004; **see Extended paper section 1.1 to 1.1.4 for further discussion regarding the challenges in dementia assessment and diagnosis).** This situation is further compounded by the limited progress made in developing curative or preventative treatments (Pryse-Phillips, 1999). However, Anti-cholinesterase Inhibitors (AchI), which aim to halt and reverse decline in both cognitive abilities and activities of daily living (Foy & Starr, 2000), have led to benefits, if somewhat modest, for some (Birks & Harvey, 2006; Starr & Lonie, 2008), and have generated considerable interest in recent years (Melzer, 1998; Doyle, 2001; Singh & O'Brien, 2009; **see Extended paper section 1.3 to 1.3.1 for further discussion regarding the evidence base for AchI).**

It is estimated that 24.3 million people around the world have dementia, with higher numbers in developed countries (Ferri *et al.,* 2005). Of the 683,597 people with dementia in the UK, 11,392 are from Black and Minority Ethnic (BME) backgrounds (Albanese *et al.,* 2007). However these could be underestimations due to scarce epidemiological studies in developing countries (Ferri *et al.,* 2007), the lack of culturally appropriate cognitive assessments (Chen, 2004), and differences in cultural beliefs about ageing (Ganguli *et al.,* 1995; Lin & Lee, 1997; Rait *et al.,* 2000; Richards *et al.,* 2000; Butt & O'Neil, 2004; Purandare *et al.,* 2007; Ramsey *et al.,* 2009; **see**

Extended paper section 1.2 to 1.2.4 for further discussion about international and UK specific prevalence estimates and challenges to establishing prevalence estimates in developing countries and BME groups within western countries).

Most BME elders migrated to the UK as young adults, and relatively few have gained fluency and literacy in English (Lindesay *et al.,* 1997; Parker & Philip, 2006). The most commonly used cognitive measures have been validated only among English speaking Caucasians, and require fluency in English and spoken language to be completed (Lampley-Dallas, 2001; Manly & Jacobs, 2001). Hence, there is a need for cross-cultural neuropsychological assessments, which has been repeatedly emphasised (Maj *et al.,* 1993; Nell, 1999; Chesters, 2007; Manly & Echemendia, 2007; **see Extended paper section 1.6 and 1.7 for further discussion about cognitive assessment with minority groups).**

The National Institute for Health and Clinical Excellence (NICE) decided to restrict the availability of AchI to patients with Alzheimer's Disease (AD) of moderate severity (NICE, 2007). This corresponds to a score of 10-20 points on the Mini-Mental State Examination (MMSE, Folstein *et al.,* 1975). This decision was met with widespread controversy and disagreement (Pelosi *et al.,* 2006; Rodda & Walker, 2009). The debate is over the magnitude of the effect and whether benefits are cost-effective for the state (Singh & O'Brien, 2009; **see Extended paper section 1.4 to 1.4.1 for further discussion**). Furthermore, despite being the most frequently used and highly cited assessment for screening cognitive functioning in older people (Tombaugh & McIntyre, 1992; MacKenzie *et al.,* 1996; Davey & Jamieson, 2004; Nilsson, 2007; Mitchell, 2009), the MMSE is confounded by multiple sources of score variance resulting in its psychometric instability, which is well documented (Brayne &

Calloway, 1990; Strauss *et al*., 2006 - **see Extended paper section 1.5 for critique of the MMSE).**

A High Court judge ruled that NICE had acted illegally by not providing specific guidance for groups for whom the MMSE was not administrable, such as those with language impairments or for whom English is not a first language (Syrett, 2007). The lack of progress made in developing alternative objective assessment procedures for these groups has resulted in inequalities in assessment practices and limits clinician's capacity to deliver evidenced based treatments **(see Extended paper section 1.4.2 for more detailed discussion).** BME and non-English speakers are also under-represented in Randomised Controlled Trials (RCT) for AchI and dementia research in general (Hussain-Gamble *et al.,* 2004; Lopez *et al.,* 2008 – **see extended paper section 1.3.2 for further information about exclusion of BME groups in research**).

Subsequently, the aim of this study was to explore whether alternative objective cognitive tests, which did not rely on spoken language from participants, could provide equivalent information to the moderate range on the MMSE. In addition, to explore whether optimum cut-off scores could be identified on these cognitive tests, which where equivalent to the cut-offs on the MMSE that infer eligibility for treatment with AchI.

METHODS

2.1 Design

This study employed a cross-sectional design in which all participants completed six index tests as well as the MMSE as the reference standard (described in the next section).

2.2 Sample

The sample consisted of 20 participants aged between 65 and 90 years. The initial sample size was n=30. Six subjects were excluded because they did not have capacity to consent; two because they withdrew consent; one because a carer did not consent and finally one because they were presenting with high levels of anxiety. Thus, the sample size reduced to n= 20. To be included participants needed to be >65 years, have a diagnosis of a dementia type illness, or were presenting with age related cognitive difficulties, and were fluent in English. Participants were excluded if they had a mood disorder, lacked capacity to consent, had fine motor difficulties, or if they had visual, or hearing impairments which would have prevented them from completing the cognitive tests **(see Extended paper section 2.3 for further details about the exclusion criteria).**

2.3 Procedures

Participants were recruited with the approval of the regional ethics committee and Research and Development approval from two NHS sites in the UK. Participants who met the inclusion criteria were referred by Health Care Professionals (HPCs) working within older people's Mental Health Services. Initial appointments were arranged via telephone. All measures were administered during this appointment, taking approximately 1 to 1.5 hours. Tests were administered in participant's homes or in day hospital settings. All tests were administered using standardised procedures outlined in respective published manuals **(see Extended paper Appendix 13 for flow diagram of recruitment, and section 2.5.1 and 2.5.2 for more detailed information about recruitment and data collection).**

Measures

*The University of San Diego Brief Assessment of Capacity to Consent (*UBACC; Jeste *et al.,* 2007*):* assessed using ten items assessing participants' understanding and appreciation of the study information. Those unable to answer questions with sufficient detail were excluded **(see Extended paper section 2.4.1 for further information about this measure; extended paper Appendix 7).**

Demographics Questionnaire

Descriptive information about the sample, such as age, gender, education, ethnicity, relationship status, diagnosis, medication, history of head injury and accommodation were collected from the participant or carer **(see Extended paper Appendix 10)**.

*MMSE (*Folstein *et al.,* 1975; Folstein, *et al.*, 2000*)*: consisted of 11 items assessing orientation to time, and place, registration and recall, concentration, naming objects, repeating a phrase, following verbal and written directions, writing and constructional praxis. It took 15 minutes to complete **(see Extended paper section 2.4.3 for further details on the administration and psychometric properties for this test).**

Index tests

Instruments were selected on the basis of their value in assessing cognitive domains, which are affected by dementia, with sufficient reliability. They did not rely on participants' spoken language, requiring either drawing or pointing to communicate their responses instead. All instructions were delivered verbatim in English by the first author. No studies have explored the predictive utility of the selected measures to identify positive and negative cases for treatment with AchI. The index tests were:

*Rey Complex Figure Test (*RCFT; Meyers & Meyers, 1995*)*: consisted of four items assessing visual and spatial constructional ability, visual speed of processing, immediate, delayed and recognition memory, in the visual modality only. This involved copying a figure whilst being timed, then drawing it from memory following a 3, then 30 minute delay. This test took 15-20 minutes to complete (minus 30 min delay; **see Extended paper section 2.4.4 for further details on the administration and psychometric properties for this test).**

*Ravens Colour Progressive Matrices (*RCPM; Raven *et al.,* 2003*):* assessed fluid intelligence through reasoning ability and is a frequently used test in gerontological research (Cockburn & Smith, 1991; Smits *et al.,* 1997). The test took a maximum of 25 minutes to complete **- see Extended paper section 2.4.5 for further details on the administration and psychometric properties for this test).**

*Symbol Digit Modality Test (*SDMT; Smith, 1991*):* assessed attention, visual scanning, concentration, and motor and psychomotor speed. Participants were instructed to fill in as many blank spaces under a nonsense symbol with the correct matching numbers according to a code. They were given 90 seconds to fill as many gaps as possible, and this took 5 minutes to complete **-see Extended paper section 2.4.6 for further details on the administration and psychometric properties for this test).**

*Brixton Test (*Burgess & Shallice, 1997*):* assessed behavioural regulation, or concept (or 'rule') attainment ability, which is related to executive functioning. Deficits in executive functioning are a notable area of cognitive decline in dementia (Brennan *et al.*, 1997; Rainville, 2002; Bielak *et al.*, 2006). This test took approximately 15-20 minutes to complete in this sample **(see Extended paper section**

2.4.7 for further details on the administration and psychometric properties for this test).

*Clock Drawing Test (*CDT*;* Tuokko *et al.,* 1995*)*: assessed visuospatial, constructional and executive difficulties. Participants were instructed to draw the numbers in a pre-drawn clock face and then asked to draw in the hands to read ten past eleven. The test took five minutes to complete **(see Extended paper section 2.4.8 for further details on the administration and psychometric properties for this test).**

*Colour Trails Test (*CTT; D'Elia *et al.,* 1996*)*: assessed speed for visual scanning, divided attention, cognitive flexibility, motor function and speed of information processing. It has been used to assess cognitive functioning in cross-cultural settings (Maj *et al.*, 1993; Dugbartey, 2000; Elkin-Franston *et al.*, 2007), and can be completed in 5- 10 minutes **(see Extended paper section 2.4.9 for further details on the administration and psychometric properties for this test).**

Statistical analysis was conducted using correlation and Receiver Operating Characteristic (ROC) curve analysis (Altman & Bland, 1994; Swets, 1988). Data was analysed using SPSS for windows (Version 16).

RESULTS

Characteristics of the sample

The initial sample size was n=30. Six subjects were excluded because they did not have capacity to consent; two because they withdrew consent; one because a carer did not consent and finally one because they were presenting with high levels of anxiety. Thus, the sample size reduced to n= 20. Eleven (55%) males and nine (45%) females took part in this study. The mean age of participants was 77.6 years

of age (SD=7.2), ranging from 65 to 90 years. The mean years of education was 10.7 years (SD = 3.3), ranging from 9 to 23 years. All 20 participants were living in their own homes. Eighteen (90%) of the participants classified themselves as White British, one as White Irish (5%) and one (5%) as White Scottish.

Fifteen subjects (75%) were diagnosed with AD, three (15%) were diagnosed with mixed AD and Vascular Dementia, and two were presenting with mild cognitive impairment. The mean score on the MMSE was 22 (SD = 4.4), ranging from 14 to 29 points. Only three participants who scored within the severe range on the MMSE (0-9) were referred to this study and these three were unable to provide informed consent. Hence, it was not possible to explore the 0-9 cut off.

There were 8 (40%) positive cases (scored <20 on MMSE, eligible for AchI) and 12 (60%) negative cases (scored >21 on the MMSE, not eligible for AchI) in this sample. In the 8 positive cases the score range was 14 to 20 points on the MMSE, and all were being treated with an AchI. In the twelve negative cases, the score range was 21 to 29 points on the MMSE, and ten (83%) were being treated with an AchI. In total 18 (90%) participants were being treated with an AchI (n=16, 80%, with Aricept; n=2, 5%, Exelon).

Following normality analysis, Pearson product moment coefficient was calculated between scores on the MMSE and scores on the RCFT-visual construction, RCFT-speed of process, RCFT-recognition memory, RCPM, SDMT, Brixton test, CDT and CTT2. Spearman correlation coefficient was calculated between scores on the MMSE and the RCFT – immediate memory, RCFT-delayed memory and CTT1 **(see Extended paper section 3.4 to 3.5 for details regarding exploration of distribution of scores).** The correlations obtained are shown in Table 1.

Table 1: Pearson and Spearman correlation coefficients for MMSE and index tests

Index tests‡	N	MMSE	Significance
RCFT VC	19	r= .609**	.006
RCFT SP	19	.128	.602
RCFT I	19	.131	.594
RCFT D	19	.370	.119
RCFT R	19	r= .496*	.031
RCPM	20	r= .452*	.045
SDMT	20	r= .670**	.001
Brixton Test	18	-.403	.098
CDT	20	r= -.577**	.008
CTT1	18	r= -.576*	.012
CTT2	12	-.372	.234

‡ RCFT VC, Rey Complex Figure Test Visual Construction Trial; RCFT SP, Rey Complex Figure Test Speed of Processing; RCFT I, Rey Complex Figure Test Immediate Memory; RCFT D, Rey Complex Figure Test Delayed Memory; RCFT R, Rey Complex Figure Test Recognition Memory; RCPM, Ravens Colour Progressive Matrices; SDMT, Symbol Digit Modalities Test; CDT, Clock Drawing Test; CTT1 & 2, Colour Trails Test 1 & 2.

** = Correlation is significant at p< 0.01 level, 2-tailed

* = Correlation is significant at p< 0.05 level, 2-tailed

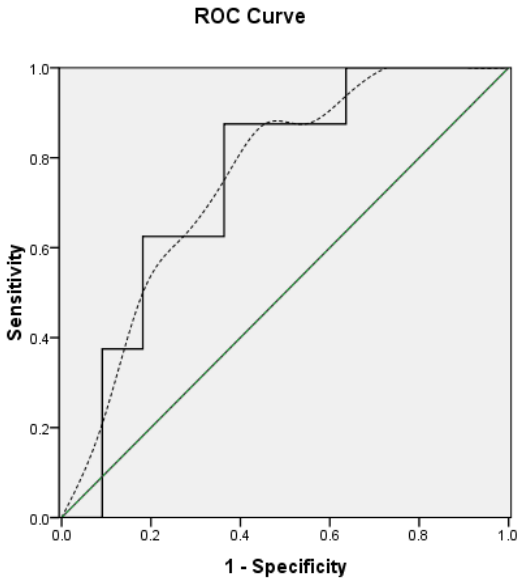
ROC analysis was carried out with measures that significantly correlated with the MMSE. ROC curves were plotted to show the trade off between sensitivity and specificity for all scores, allowing for the choice of most appropriate cut-off. Particular regions of the ROC plane correspond with particular type of diagnostic performance. The extreme lower left corner represents a measure that classifies all participants as not being eligible for AchI. The extreme top right corner represents a measure that classifies all participants as eligible for AchI. The diagonal line crossing through the plane represents chance. Performance below this line is worse than chance. The ROC curves are presented in Fig 1, 2 and 3 **(see Extended paper section 3.9 to 3.9.3 for more details about ROC analysis**).

The area under the ROC curve values were as follows: RCFT visual construction (0.750, 95%, CI .524 - .976), RCFT recognition memory (0.801, 95%, CI .590 – 1.012), RCPM (0.573, 95%, CI 0.298 – 0.848), SDMT (0.708, 95%, CI 0.469 – 0.947), CTT1 (0.818, 95%, CI 0.610 – 1.027) and the CDT (0.734, 95%, CI 0.479 – 0.990). Cut-off scores with adequate sensitivity and specificity were identified for all the above measure apart from the RCPM, which had predictive ability that was equal to chance. The sensitivity, specificity, Positive Predictive Value, Negative Predictive Value and overall Discriminant Ability (DA) for chosen cut-offs were calculated using cross tabulations and kappa analysis. The results are shown in Table 2 **(see Extended paper section 3.10 to 3.15 for a detailed description of how ROC analysis was carried out for each measure that significantly correlated with the MMSE)**.

Table 2: Optimum cut-off scores on index test which are equivalent to the cut-off (<20) on the MMSE

Index tests *	**Cut-off score**	**Sen %**	**Spec %**	**PPV %**	**NPV %**	**Discriminant ability %**
RCFT VC	<20.5[a]	87	64	64	87	75
RCFT R	<14.5[a]	87	72	70	89	79.5
SDMT	<11[a]	75	66	60	80	70.8
CTT1	>144.5 seconds[b]	86	64	60	88	75
CDT	>9[b]	75	75	60	80	71
	* RCFT VC, Rey Complex Figure Test Visual Construction; RCFT R, Rey Complex Figure Test Recognition; SDMT, Symbol Digit Modalities Test; CTT1, Colour Trails Test 1; CDT, Clock Drawing Test; Sen, Sensitivity; Spec, Specificity; PPV, Positive Predictive Value, and NPV, Negative Predictive Value. [a] lower score indicated greater impairment [b] Higher score indicated greater impairment					

RCFT- Visual Construction



RCFT- Recognition Memory

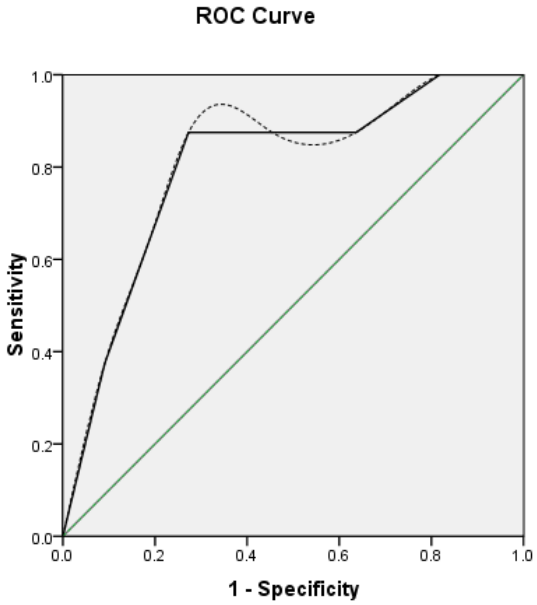
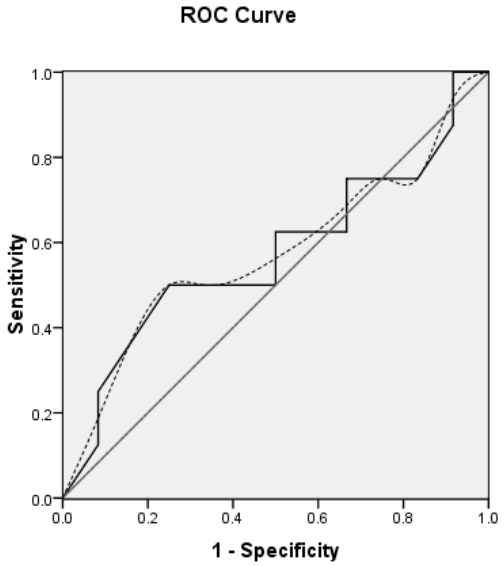


Fig1. ROC curves for RCFT visual construction and recognition memory subtests

Ravens Colour Progressive Matrices



Symbol Digit Modalities Test

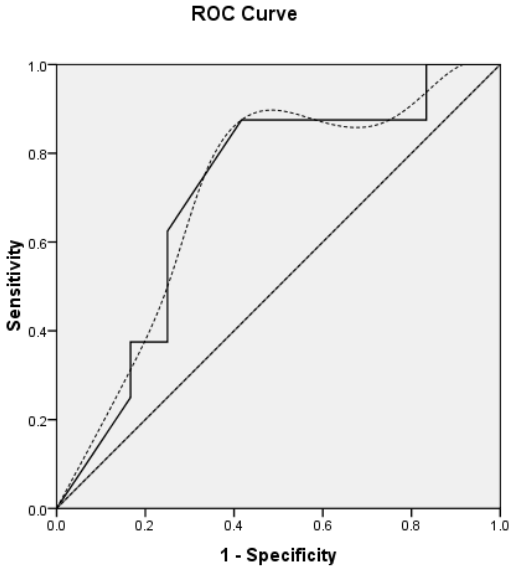
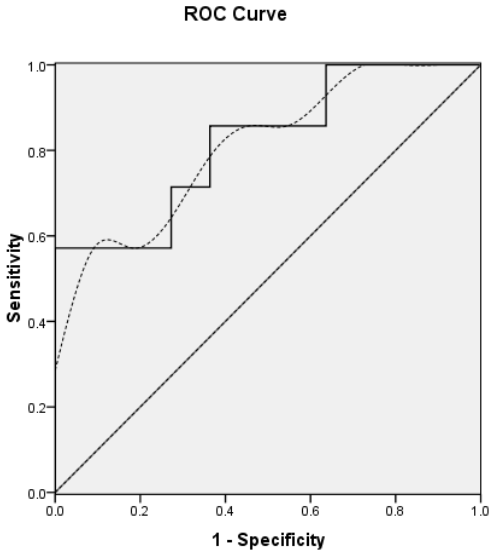


Fig2. ROC curves for RCPM and SDMT

Clock Drawing Test



Colour Trails Test 1

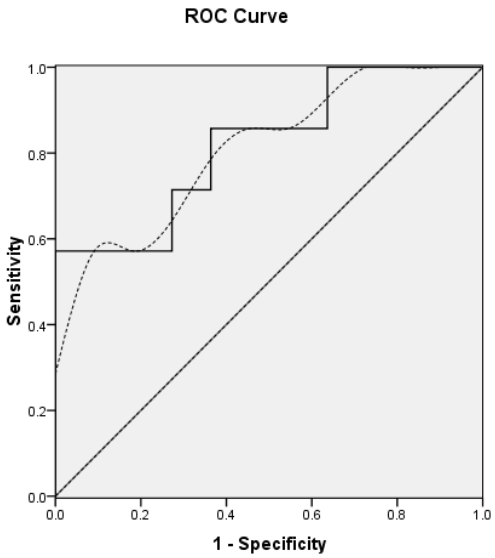


Fig3. ROC curves for CDT and CTT1

DISCUSSION

The aim was to explore the utility of alternative cognitive tests, which do not rely on spoken language, to address the lack of objective cognitive assessments for people who are not fluent in English and those who are non-verbal, to decide if they are eligible for treatment with an AchI (NICE, 2007).

Relationship between MMSE and index tests

High correlation coefficients were found for five of the six index tests (see Table 1). Out of 11 items in total, the MMSE had high statistically significant correlations with six items assessing: visual construction (RCFT); recognition memory (RCFT), reasoning ability (RCPM); attention, visual scanning, motor speed (SDMT); visual construction and executive difficulties (CDT); attention, speed for visual scanning, divided attention, cognitive flexibility, motor function and speed of information processing (CTT1). In line with previous studies we found a high correlation between the MMSE and the CDT (Mendez *et al.,* 1992; Schramm *et al.,* 2002). We did not find a statistically significant correlation between the MMSE and specific tests of executive functioning, such as the Brixton test and CTT2, which is a key limitation of the MMSE as a dementia screen. However, the Brixton test is confounded by short term memory difficulties (Rainville *et al*., 2002). Participants reported forgetting of previous items, which was not surprising due to the 55 trials. Furthermore, participants' reported difficulty comprehending instructions for CTT2, which led to missing data from eight participants. However, difficulties with the CTT2 could also be explained by order effects, as it was the last measure to be administered, thus performance may have been confounded by fatigue **(see extended paper section 4.1 for further discussion related to the utility of index tests).**

The high correlations with the other common tests of cognitive functioning were a somewhat surprising finding given that the MMSE

has previously had weak coefficients with cognitive tests measuring similar constructs (Strauss *et al.,* 2006). These findings provide partial evidence that as a screening tool the MMSE performs quite well (Tombaugh & McIntyre, 1992; Han *et al.,* 2000; Strauss *et al.,* 2006). However, a limitation of this study is that a clinical sample, which was being treated with an AchI, was used. These two variables could have been confounding variables. Therefore, it is unclear whether the correlations between the MMSE and the index tests were due to a true relationship between the domains that were being measured, or due to the effect of the medication or diagnosis. This could be rectified in a future study, where these two variables are controlled and a non-clinical sample is used **(see extended paper section 4.4 for further discussion regarding the limitations of this study)**.

Predictive accuracy of index tests & cut off scores

All measures, apart from the RCPM, obtained AUC values above .708, which demonstrates moderate accuracy (Bewick *et al.,* 2004; Linden, 2006). A previous study has obtained the same finding in relation to the RCPM (Blake *et al.,* 2002). AUC values were highest for RCFT recognition memory (.801) and the CTT1 (.818), which is moderate to high in predictive accuracy (Fischer *et al.,* 2003). The SDMT had the lowest AUC value (.709).

It was possible to identify cut-off scores which were equivalent to the <21 point cut off on the MMSE. Therefore, clinicians would be able to use these cut offs to aid decision making about prescribing AchI. An optimum cut-off of <20.5 for the RCFT visual construction, had good sensitivity (87%) and moderate specificity (64%). Therefore, 87% of participants scored 20.5 or below on this measure, which was the indicator for eligibility for treatment with AchI, and 64% scored above this cut-off. However the usefulness of a test is judged by whether or not those who tested positive using this cut-off were classified

correctly (Altman & Bland, 1994; Linden, 2006). For this test, the PPV was moderate (64%) and the NPV (87%) was good. This suggested that only 64%, who were below the cut-off on the RCFT visual construction trial, actually had a MMSE score within the moderate range, and 87% of those who scored above the cut-off on the RCFT, also obtained above the cut-off on the MMSE. A similar pattern was observed for the four other index tests (see Table 2). The PPV for all the cut-offs for the index tests were adequate. However, these results have to be viewed in light of the limitations of this study, namely the small sample size and the unequal prevalence of positive (8, 40%) and negative cases (12, 60%). This would have reduced the PPV and increased the NPV as they are sensitive to unequal prevalence rates (*Bewick et al,* 2004). In addition, it is possible that the cut-off scores for all five index tests (see Table 2) could have been different with a none clinical sample, or if there were more participants who scored within the moderately severe range (10-15). Therefore further research which manipulates these variables would enable us to learn more about how they impact on the predictive utility of the chosen measures and their cut-off scores.

The higher sensitivity and lower specificity values could reflect the high level of cognitive impairment in this sample. The cut-offs chosen were all within the impaired range for their respective measures. For measures such as the SDMT and RCFT recognition trial the cut-off score was so low that it is likely that if this study included more people who scored within the severe (0-9) or moderately severe (10-15) range on the MMSE, they may have scored zero. Hence it may be less meaningful to use these tests for these groups in clinical practice to monitor change.

Furthermore, the RCFT recognition trial had the highest PPV and NPV, however all subtests of the RCFT would need to be administered in

clinical practice to obtain the recognition score. Any alternative test to the MMSE would need to balance high predictive accuracy with the preference for it being quick and require minimal training (Glasser, 1993; Mitchell, 2009). Therefore, the SDMT, CDT, CTT1 and RCFT visual construction may be adequate alternatives to the MMSE in terms of time.

The implications of these results are potentially considerable for GPs, older people's Mental Health Services and memory clinicians. This is because clinicians carry out cognitive assessments with people with dementia routinely. However, the findings of this study are only preliminary and further research needs to be carried out before the index tests can be used to make decisions about eligibility for treatment with AchI **(see extended paper section 4.3 for further discussion regarding the clinical implications of this research)**.

In relation to clinical practice, this study also found that in this sample 80% who were not eligible for AchI, according to their MMSE score, were being treated with an AchI. Therefore, providing partial evidence that clinicians may not be adhering strictly to guidelines (NICE, 2007). Further research exploring treatment practices is needed to establish how prevalent this is across services and also the evidence base that is informing such decisions **(See Extended paper section 4.2 for further discussion related to the prescription of AchI).**

Finally, it is important to bear in mind that this research was based on the assumption that the MMSE is a gold standard assessment. There is limited empirical evidence to support that the 'moderate' range on the MMSE is a valid construct on which to base decisions about treatment eligibility (Olin & Zelinsky, 1991). This was going beyond its intended use as a dementia screen (Folstein *et al.,* 1975). Further research about the neurological, neuropsychological and

functional parameters of the 'moderate' range is warranted, if the guidelines are to remain valid (Cerejeire & Mukaetova-Ladinska, 2007; **see extended paper section 4.6 for further discussion about expanding on this research).**

Four key points

1. Cognitive tests that do not require spoken language may have utility in assessing for eligibility for treatment with AchI.
2. Research exploring the cross cultural equivalence of these measures is necessary.
3. The clinical parameters of the "moderate" range remain unclear.
4. AchI are prescribed to people who score above the cut-off on the MMSE.

CONFLICT OF INTEREST
None Known

ETHICAL APPROVAL
The study was approved by the regional ethics committee and the NHS.

REFERENCES

Albanese, E., Banerjee, S., Dhanasiri, et al. 2007. Dementia UK: report to the Alzhiemer's society, Kings College London and London school of economics and political science. Alzheimer's Society.

Altman, D.G. & Bland, M. 1994. Diagnostic tests 2: Predictive values. BMJ, **309**, 102.

Bewich, V., Cheek, L. & Ball, J. 2004. Statistics review 13: Receiver operating characteristic curves. Crit care Med, **8** (6), 508-512.

Bielak, A.A.M., Mansueti, L., Strauss, E. & Dixon, R.A. 2006. Performance on the Hayling and Brixton tests in older adults: norms and correlates. Arch Clin Neuropsychology, **21**, 141-149.

Birks, J. & Harvey, R.J. 2006. Donepezil for dementia due to Alzheimer's disease (Review). The Cochrane Database of Systematic Reviews, [Online], **4**, Available at: http://mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD001190/frame.html [Accessed 10th October 2007].

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

Brayne, C. & Calloway, P. 1990. The association of education and socioeconomic status with the mini-mental state examination and the clinical diagnosis of dementia in elderly people. Age Ageing, **19**, 91-96.

Brennan, M., Welsh, M.C., & Fisher, C.B. 1997. Ageing and Executive function skills: An examination of a community-dwelling older adult population. Percept Mot Skills, **84**, 1187-1197.

Burgess, P. W., & Shallice, T. 1997. The Hayling and Brixton tests. Thurston, Suffolk: Thames valley Test Company.

Butt, J., & O'Neil, A. 2004. 'Lets Move On': Black and Minority Ethnic Older People's Views on Research Findings. Joseph Rowntree Foundation, [Online], **175**, (1). Available at: www.jrf.org.uk/knowledge/findings/socialcare/564.asp [Accessed December 2007].

Cairns, R., Evans, J. & Prince, M. 2004. The impact of NICE guidelines on the diagnosis and treatment of Alzheimer's Disease among general medical hospital inpatients. Int J of Geriatr Psychiatry, **19**, 800-802.

Cerejeira, J. & Mukaetova-Ladinska, E.B. 2007. Treating dementia: will the NICE guidance 2006 change our practice? Age Ageing, **36**, 605-606.

Chen, C.P.L.-H. 2004. Transcultural expression of subcortical vascular disease. J Neurol Sci, **226**, 45-47.

Chesters, M.J. 2007. Cross-cultural and cross-linguistic neuropsychology. Division of Neuropsychology Newsletter. British Psychological Society, **12**, 6-8.

Cockburn, J., & Smith, P.T. 1991. The relative influence of intelligence and age on everyday memory. J Geront, **46**, 31–36.

Davey, R.J., & Jamieson. S. 2004. The validity of using the mini mental state examination in NICE dementia guidelines. J Neurol Neurosurg Psychiatry, **75**, 343–344.

D'Elia, L.F., Satz, P., Uchiyama, C. L., & White, T. 1996. Colour Trails Test. Professional manual. Odessa, FL: Psychological Resources.

Doyle, Y. 2001. Equity in the new NHS: hard lessons from implementing a local healthcare policy on denepezil. BMJ, **323**, 222-224.

Dugbartey, A.T., Townes, B.D., Mahurin, R.K. 2000. Equivalence of the Colour Trails Test and Trail Making Test in Non-native English-Speakers. Arch Clin Neuropsychology, **15** (5), 425-431.

Elkin-Frankson, S., Lebowitz, B.K., Kapust, L.R., Hollis, A.M. & O'Conner, M.G. 2007. The use of the colour trails test in the assessment of driver competence: preliminary report of a culture fair instrument. Arch Clin Neuropsychology, **22**, 631-635.

Ferri, C.P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., et al. 2005. Global prevalence of dementia: a Delphi study. Lancet, **366**, 2112-2117.

Fischer, J.E., Bachmann, .LM., Jaeschke, R. 2003. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intens Care Med, **29**, 1043–1051.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. 1975. Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res, **12**, 189-198.

Folstein, M.F., Folstein, S.E. & Fanjiang, G. 2000. Mini-Mentla State Examination: clinical guide. Florida, USA: Psychological Assessment Resources, Inc.

Foy, J.M. & Starr, J.M. 2000. Assessment and treatment of dementia in medical patients. Psychother Psychosom, **69**, 59-69.

Galasko, D, Klauber, M. R., Hofstetter, C. R., Salmon, D. P., Lasker, B., & Thai, L. J. (1990). The Mini-Mental State Examination in the early diagnosis of Alzheimer's disease. Arch Neurology, **47**, 49-52.

Ganguli, M., Ratcliff, G., Chandra, V., Sharma, S., Gilby, J., Pandav, R., Belle, S., Ryan, C., Baker, C., Seaberg, E. & Dekosky, S. 1995. A Hindi version of the MMSE: The development of a cognitive screening instrument for a largely illiterate rural elderly population in India. Int J Geriatr Psychiatry, **10**, 367-377.

Glasser, M. (1993). Alzheimer's disease and dementing disorders: practices and experiences of rural physicians. Am J Alzheimer's Dis Other Demen, **8** (4), 28-35.

Green, J. 2000. Neuropsychological Evaluation of the Older Adult: A Clinician's Guidebook. London, UK: Academic press.

Han L, Cole M, Bellavance F, McCusker J, Primeau F. 2000. Tracking cognitive decline in Alzheimer's disease using the mini-mental state examination: a meta-analysis. Int Psychoger, **12** (2), 231–47.

Hussain-Gambles, M., Leese, B., Atkin, K., Brown, J., Mason, S. & Tovery, P. 2004. Involving south Asian patients in clinical trials. Health Technol Assess, **8** (42).

Jeste, D.V., Palmer, B.W., Appelbaum, P.S., Golshan, S., Glorioso, D., Dunn, L.B et al. 2007. A brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry, **64** (8), 966-974.

Lampley-Dallas, V.T. 2001. Neuropsychological screening tests in African Americans. J Natl Med Assoc, **93** (9), 323–328.

Lin. H.C. & Lee, E. 1997. Epidemiology of dementia among the Chinese. Neurol J South East Asia, **2**,129 – 34.

Linden, A. 2006. Measuring diagnostic and predictive accuracy in disease management: anintroduction to receiver operating characteristic (ROC) analysis. J Eval Clin Pract, **12** (2), 132-139.

Lindesay, J., Jagger, C., Mlynik-Szmid, A., Sinorwala, A., Peet, S. & Moledine, F. 1997. The Mini-Mental State Examination (MMSE) in an elderly immigrant Gujarati population in the United Kingdom. Int J Geriatr Psychiatry, **12** (12), 1155-1167.

Lopez, O.L., Mackell, J.A., Sun, Y., Kassalow, L.M., Xu, Y., McRae, T. & Li, H. 2008. Effectiveness and safety of donepezil in Hispanic patients with Alzheimer's Disease: a 12-week open-label study. J Natl Med Assoc, **100** (11), 1350-1358.

MacKenzie, D.M., Copp, P., Shaw, R.J. & Goodwin, G.M. 1996. Brief cognitive screening for the elderly: a comparison of the mini-mental state examination (MMSE), abbreviated mental test (AMT) and mental status questionnaire (MSQ). Psychol Med, **26**, 427-430.

Maj, M., D'Elia, L., Satz, P., Janssen, R., Zaudig, M., Uchiyama, C., Starace, F., Galderisi, S & Chervinsky, A. 1993. Evaluation of Two New Neuropsychological Tests Designed to Minimize Cultural Bias in the Assessment of HIV-1 Seropositive Persons: A WHO Study. Arch Clin Neuropsychology, **8**, 123-135.

Manly, J.J., & Jacobs, D.M. 2001. Future directions in neuro-psychological assessment with African Americans. In Ferraro, F.R 2001. Minority and cross cultural aspects of neuropsychological assessment. Lisse, Netherlands: Swets and Zeitlinger, 79-96.

Manly, J.J & Echemendia, R.J. 2007. Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. Arch Clin Neuropsychology, **22**, 319-325.

Mendez, M.F., Ala, T. & Underwood, K.L 1992. Development of scoring criteria for the clock-drawing task in Alzheimer's disease. J Am Geriatr Soc, **40**, 1095-1099.

Meyers, J., & Meyers, K. 1995. Rey Complex Figure and Recognition Trial: Professional manual. Odessa, FL: Psychological Assessment Resources.

Mitchell, A.J. 2009. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. J Psychiatr Res, **43**, 411-431.

National Institute for Clinical Excellence. 2007. Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease (amended). Available at: http://www.nice.org.uk/nicemedia/pdf/TA111fullversionamendedSept07.pdf [Accessed 30th September 2007].

Nell, V. 1999. Cross-cultural neuropsychological assessment: theory and practice. New York: Lawrence Erlbaum Associates.

Nilsson, F.M. 2007. Mini-mental state examination (MMSE) - probably one of the most cited papers in health science. Acta Psychiatrica Scandinavica, **116** (2), 156-7.

Olin, J.T. & Zelinski, E.M. (1991). The 12-month reliability of the mini-mental state examination. J Consult Clin Psychol, **3** (3), 427-432.

Parker, C. & Philip, I. 2006. Screening for cognitive impairment among older people in black and minority ethnic groups. Age Ageing, **33** (5), 447-452.

Pelosi, A.J., McNulty, S.V., & Jackson, G.A. 2006. Role of cholinesterase inhibitors in dementia care needs rethinking. BMJ ,**333**, 491–493.

Pryse-Phillips, W. 1999. Do we have drugs for dementia? Archives of Neurology, **56**, 735-737.

Purandare, N., Luthra, V., Swarbrick, C. & Burns, A. 2007. Knowledge of dementia among south Asian (Indian) older people in Manchester UK. Int J Geriatr Psychiatry, **22**, 777-781.

Rainville, C., Amieva, H., Lafont, S., Dartigues, J., Orgogozo, J-M. & Fabrigoule, C. 2002. Executive function deficits in patients with dementia of the Alzheimer's type A study with a tower of London task. Arch Clinl neuropsychology, **17,** 513-530.

Rait, G., Burns, A., Baldwin, R, Morley, M., Chew-Graham & St Leger, S. 2000. Validating screening instruments for cognitive impairment in older south Asians in the United Kingdom. Int J Geriatr Psychiatry, **15**, 54-62.

Ramsey, V., Stevens, S., Bryan, K., Binder, J. & Cockle-Hearne, J. 2009. Using the Barnes language assessment with with older ethnic minority groups. Int J Geriatr Psychiatry, **24**, 426-431.

Raven, J., Raven, J.C., & Court, J.H. 2003. Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview. San Antonio, TX: Harcourt Assessment.

Richards, M., Brayne, C., Dening, T., et al. (2000). Cognitive function in UK community-dwelling African Caribbean and White Elders: a pilot study. *International Journal of Geriatric Psychiatry,* **15***,* 621–630.

Robillard, A. 2007. Clinical diagnosis of dementia. Alzheimer's Dementia, **3**, 292-298.

Rockwood, K., Bouchard, R.W., Camicioli, R. & Leger, G. 2007. Toward a revision of criteria for the dementias. Alzheimer's & Dementia, **3**, 428-440.

Rodda, J. & Walker, Z. 2009. Ten years of cholinesterase inhibitors. Int J geriatr Psychiatry, **24**, 437-442.

Schramm, U., Berger, G., Muller, R., Kratzsch, T., Juergen, P., & Frolick, L. 2002. Psychometric properties of the clock drawing test and MMSE or short performance test (SKT) in dementia screening in a memory clinic population. Int J Geriatr Psychiatry, **17**, 254-260.

Singh, B. & O'Brien, J.T. 2009. When should drug treatments be started for people with dementia? Maturitas, **62**, 230-234.

Smith, A. 1991. Symbol digit modalities test. Los Angeles, CA: Western Psychological services.

Starr, J.M. & Lonie, J. 2008. Estimated pre-morbid IQ effects on cognitive and functional outcomes in Alzheimer's Disease: a longitudinal study in a treated cohort. BMC Psychiatry. Available at: http://www.medscape.com/viewarticle/575011_print [Accessed December 2007].

Strauss, E., Sherman, E.M.S. & Spreen, O. 2006. A compendium of neuropsychological tests: administration, norms and commentary (3rd ed). New York, USA; Oxford University Press.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240,* 1285-1293.

Syrett, K. 2007. NICE and judicial review: enforcing accountability for reasonableness through the courts? Eisai limited v. National institute of clinical excellence. Medical Law Review. 1-14.

Tombaugh, T. & McIntyre, N. 1992. The Mini-Mental State Examination: A comprehensive review. J Am Geriatr Society, **40**, 922-935.

Tuokko. H., Hadjistavropoulos, T., Miller, A., Horton, A. & Beattie, B.L. 1995. The clock test: administration and scoring manual. New York, USA: Multi-Health Systems Inc.

1. Extended Introduction

Chapter Overview

This section will present additional background literature related to the aetiology, assessment, diagnosis and drug treatments for dementia type illnesses. The literature will firstly be reviewed to describe the evidence-base currently informing clinical decision making, and how this can be applied to people who do not speak English (such as ethnic minority groups). The implications of the National Institute for Health and Clinical Excellence (NICE) guidance (NICE, 2007) for the prescription of Anti-cholinesterase Inhibitors (AchI) for people with Alzheimer's Disease (AD), and the role of the Mini-Mental State Examination (MMSE; Folstein, Folstein & McHugh, 1975) within these, will be critically discussed. The objective here being to highlight the limitations of this guidance for older people in general and then more specifically for older people who have limited proficiency in spoken English. Finally, literature will be drawn from the field of cross-cultural neuropsychology to consider the potential of alternative cognitive assessments for people who do not speak English.

This paper will focus on the most common forms of dementia, namely AD and Vascular dementia (VaD). This is because AchIs have been approved for treatment of AD. Although they have not been approved by NICE to treat Vascular Dementia (VaD) it is common for AD and VaD to be present at the same time, and if AD is assessed as being the dominant condition of the two, AchI can be prescribed (NICE, 2007).

1.1 What is dementia?

Dementia is a progressive decline in cognitive or behavioural functioning from a prior level of functioning, which is associated within decline in one or more functions, such as: short and long term

memory, abstract thinking, executive processes, language, attention and concentration, visual processing and personality changes. The progressive decline can be gradual or step-wise and eventually is severe enough to interfere with day-to-day functioning, employment or relationships (American Psychiatric Association, 2000). It is an incurable condition of major significance due to the worlds ageing population. Deterioration is associated with a continuous change in patient and care-giving problems and needs. Thus, it is imperative that ongoing efforts are made to carry out research with the aim of improving existing National Health Service (NHS) guidelines in the areas of assessment and care management that are consistent and also applicable to minority populations (National Audit Office, 2007).

1.1.1 Classification of dementia

The diagnosis of dementia is based on a careful case history. Laboratory and imaging tests as yet only provide supportive evidence (Rockwood, Bouchard, Camicioli & Leger, 2007). According to Stuss and Levine (1996) there are four ways to define and classify the dementias, 1) histologic/ neuropathologic; 2) causal risk factors (e.g. vascular, genetics, infections etc.); 3) brain region involved and 4) clinical profile (e.g. early onset, step-wise progression). To construct a differential diagnosis of the cause of dementia, the starting point is typically the clinical presentation. The usual question is whether the presentation conforms to the usual pattern seen with each dementia. However, the process of diagnosis can be confounded by many other factors such as health related co-morbidities and limitations of assessment methods for particular populations, such as people who do not speak English. Therefore a thorough clinical history, examination and longitudinal follow-up are important (Rockwood et al. 2007).

Having an indication of the pattern of progression can also give clinicians some capacity to predict the likely pattern of decline and decide upon what interventions may be most suitable. However, in later stages of the disease there is little difference in the symptoms experienced; hence diagnosis tends to become less meaningful for the individual and more difficult to differentiate for clinicians. Furthermore, even after completing assessment in all of the key areas identified by Stuss and Levine (1996) it becomes apparent that diagnosis is not straightforward. Different types of comorbidity and pathology can exist simultaneously, and the clinical decision must be made judiciously to establish the predominant pathology contributing to the dementia syndrome (Cerejeira & Mukaetova-Ladinska, 2007).

The first case of pre-senile dementia was reported in 1906 by Alois Alzheimer, a German Psychiatrist and Neuropathologist (Whitehouse, Mauer & Ballenger, 2000). Alzheimer used histology (the microscopic examination of the anatomy of diseased cells and tissues) to identify amyloid plaques, neurofibrillary tangles and arteriosclerotic changes in a female patient. This marked the first case of dementia and these early findings are consistent with histological examinations today. Until the 1960s differential diagnosis of memory difficulties within the older adult population could function with some certainty, especially with differentiating age-related benign and malignant changes (especially with AD). This could be validated against autopsy, which represented the gold standard for definitive diagnosis.

However, neurofibrillary tangles and senile plaques alone do not infer a dementia type illness as they are also present in the brains of older people without dementia (Lezak, Howieson & Loring, 2004). In addition, people with apparently identical lesions can differ widely in their cognitive functioning. Furthermore, differing sets of neuropathologic criteria yield differing estimates of dementias in the

same brains. Crystal et al (1988) showed that some elderly people met neuropathologic criteria for dementia but in fact were not showing signs of dementia (Crystal, Dickson, Fuld, Masur, Scott, Mehler et al, 1988).

1.1.2 Alzheimer's Disease

There are two sets of criteria for diagnosing AD. These are specified within the DSM-IV-TR (American Psychiatric Association, 2000) and those established by the working group on the diagnosis of AD, established by the National Institute of Neurologic and Communicative Disorders and Stroke (NINDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA; McKhann, Drachman, Folstein, Katzman, Price & Stadlan, 1984). AD is characterised by a gradual onset and gradual progressive decline of memory and at least one additional cognitive domain. Diagnosis is made on the basis of exclusion of other possible aetiologies for the progressive decline.

In light of the uncertainties around diagnosis and the need to develop better defined criteria for research participants the NINDS/ADRDA distinguishes between definite, probable and possible AD. Although the NINCDS-ADRDA criteria is widely used in research and is applicable in clinical settings, when interpreted strictly they select only so-called pure cases of AD, so exclude patients with mixed disorders, which are common. Nevertheless, in comparison to other causes the diagnostic criteria has good sensitivity (average, 81%; range, 49% to 100%), at the expense of specificity (average, 70%; range, 47% to 100%).

No single marker or set of markers for reliable positive identification of AD in living patients has yet been found (Lezak et al, 2004). Examinations of brain tissue at autopsy show the accumulation of amyloid plaques and neurofibrillary tangles. The additional

neuropathological hallmarks of AD are inflammation, neuronal loss and depletion of neurochemicals such as acetylcholine (Whitehouse, Price, Struble, Clark, Coyle & Delong, 1982; Shah & Reichman, 2006).

1.1.3 Vascular Dementia

Four diagnostic classification systems are in use for VaD, these are the California criteria, the DSM-IV-TR, the NINDS-AIREN criteria and the Hachinski Ischemic Score. In studies comparing diagnostic criteria and neuropathologic findings the NINDS-AIREN, DSM-IV-TR and California criteria had very low sensitivity (average 50%; range, 20% to 80%) but higher specificity (average 87%; range, 64% to 98%). The decreased sensitivity of diagnostic criteria is due to the considerable overlap between vascular and degenerative pathology. Some vascular pathology exists in 29% to 41% of dementia cases coming to autopsy. Pure vascular pathology accounts for dementia in only 9-10% of cases, leading some to propose that pure VaD or AD is rare.

Vascular Dementia (VaD) is a decline in cognitive functioning which can result from any of a number of vascular aetiologies. Vascular disorders such as stroke, Cerebrovascular Accident (CVA) and Transient Ischemic Attacks (TIA) are caused by pathological processes within cerebrovascular circulation. This is the supply of blood and nutrients (primarily oxygen and glucose) to the brain (Lezak et al, 2004). A disruption of normal blood flow (infarction) deprives the nervous tissue of oxygen, thus creating an area of damaged or dead tissue; an infarct. There is considerable overlap in the symptoms and presentation of vascular dementia and cerebrovascular disease. Therefore, there is an ongoing debate about the symptoms necessary to differentiate vascular dementia from vascular disease. Furthermore, the term “vascular-dementia” lacks

agreed diagnostic criteria, which results in significant differences in classification (Chen, 2004). VaD progresses in a similar way to AD, in that it gets worse over time, however progression is often 'step-wise' rather than gradual, declining slowly as the person has a new stroke. Progression of VaD can be slowed by the identification of underlying risk factors such as blood pressure.

1.1.4 Neuropsychological assessment of Dementia

One of the major roles of neuropsychological assessment is to determine whether a patient is experiencing cognitive dysfunction and to contribute to identification of aetiology so treatment can be offered (Green, 2000). When all the data of a comprehensive neuropsychological assessment – patient history, background and presenting problems, qualitative observations and quantitative scores are taken together, the examiner should have a realistic understanding of the deficits and the likely causes. Repeated testing in the dementia population can signal the pattern of progression and the type of dementia they have. Repeated testing can also be used to assess the effects of drug treatments. However, neuropsychological differences between the dementias typically show up in early stages of the disease process, hence in later stages of the disease patients become neuro-psychologically indistinguishable (Green, 2000).

A distinction has been made between cortical and subcortical dementia. Subcortical dementia describes a clinical syndrome associated with dysfunction in subcortical grey matter structures including the basal ganglia, thalamus, subthalamus, brain stem and their frontal lobe projections, as contrasted with cortical structures such as the frontal and temporal lobes (Green, 2000). Aphasia, amnesia, agnosia, and apraxia have been described as more characteristic of cortical dementias, and AD is described as the typical representation of a cortical disease. Vascular dementia is seen more

as being prototypical of a subcortical disease. However, there has been considerable debate concerning the validity of this distinction, with increasing recognition that both cortical and subcortical changes occur in most dementia disorders (Whitehouse, 1986).

In comparison to AD memory impairment in VaD may be more subtle and largely dominated by other cognitive dysfunctions (Sellal, Woltt & Marescaux, 2004). The cognitive pattern changes according to the type of VaD. In multi-infarct dementia cortical lesions may cause loss of instrumental functions manifested by aphasia, amnesia, apraxia, or agnosia. Compared to patients with AD, VaD patients typically perform better on verbal learning and memory tasks and have better delayed recall and lower rates of forgetting. Subcortical ischaemic dementia frequently impairs executive functions, attention, and speed of information processing. The clinical hallmark of VaD is dysexecutive syndrome. In addition, language is one of the first skills to be affected by dementia (Green, 2000). Therefore, the difficulty with expressive and receptive language becomes a significant barrier for communication.

1.2 Prevalence of dementia

1.2.1 International Estimates

Alzheimer's Disease International, the umbrella organisation for national Alzheimer's associations, convened an international group of experts to generate up-to-date evidence-based estimates for the prevalence and numbers of people with dementia in all regions of the world (Ferri, Prince, Brayne, Brodaty, Fratiglioni, Ganguli, et al, 2005). Twelve international experts used the Delphi method to provide prevalence estimates for every World Health Organisation (WHO) world region and the results are considered the best estimates currently available. The Delphi method involves experts systematically reviewing the evidence base independently. The

results are collated to compare estimates generated by each expert, and they then reach a consensus to provide the prevalence rates.

The numbers affected will double every 20 years to 81.1 million by 2040 (Ferri et al, 2005). These estimates were derived with the assumption that there will be no changes in mortality rates, and no preventative or curative treatments will be developed. A point of interest is that most people with dementia were noted as living in developed countries (60% in 2001, rising to 71% by 2040). It is possible that this is an under-estimation of prevalence in developing regions. Possible explanations for this are scarce epidemiological surveys in these regions (Ferri et al, 2005).There are also differences in cultural beliefs, with some traditional approaches to health and well-being conflicting with western values (Butt & O'Neil, 2004). The effects of ageing may also be experienced and treated differently in these cultures (Ramsey, Stevens, Bryan, Binder & Cockle-Hearne, 2009). For instance, some cultures remove responsibility from their elderly by taking over the older persons' responsibilities and treating them with respect. Thus the symptoms of dementia may be less obvious or tolerated as part of ageing (Rait, Burns, Baldwin, Morely, Chew-Graham & Leger, 2000; Purandare, Luthra, Swarbrick & Burns, 2007).

Finally, the lower life expectancy in some regions, such as Africa, could also account for lower rates of dementia, because people may not live into old age, therefore decreasing their potential to develop clinical conditions which have increased incidence in later life. Chen (2004) suggests that there may be methodological issues such as study design and differences in language, social customs and educational levels impacting on the validity of cognitive instruments which presume literacy, numeracy and knowledge of (and interest in) current events.

Thus, confirmation and elaboration of discrepancies in prevalence in developing countries require more prospective studies, using appropriately harmonised diagnostic procedures (de Silva & Gunatilake, 2002). However, if this difference in disease frequency is genuine, it might suggest geographical variation in the distribution of critical risk and protective factors for and against dementia (Chen, 2004).

1.2.2 Prevalence rates in the UK

In 2006, the Alzhiemer's Society commissioned the London School of Economics and the Institute of Psychiatry at Kings College London to produce a report on dementia specifically in the UK (Albanese, Banerjee, Dhanasiri, Feernandez, Ferri, Knapp et al, 2007). Ten leading UK and European experts systematically reviewed the data. Using the Delphi method, they estimated that there are 683,597 people with dementia in the UK. This number is forecast to increase to 940,110 by 2021 and 1,735,087 by 2051, an increase of 35% over the next 15 years and 154% over the next 45 years. Early onset (onset before age of 65 years) dementia is comparatively rare, accounting for 2.2% (15,034) of all people with dementia in the UK. However, this could again be an under-estimation, since these people present much later into services.

The report estimates that 416,967 people with dementia (62%) have AD in the UK. The next most common forms are Vascular dementia (VaD) and mixed dementias, accounting for 27% of all cases. However, the 'mixed dementia' classification is an anomaly, and it highlights the challenges of differential diagnosis in this area (Rockwood et al, 2007). In addition, Sellal et al (2004) state that current measures used to assess cognitive functioning place most emphasis on the domains that are typically impaired in patients with AD, therefore measures that fail to screen for executive functioning

and language deficits are likely to result in an underestimation of the incidence of VaD.

1.2.3 Older adult ethnic minority groups in the UK

The report estimates that there are 11,392 people from Black and Minority Ethnic (BME) groups with dementia, accounting for 1.7% of all people with dementia. In contrast, studies suggest that BME groups have a greater incidence of dementia, the cause attributed to unidentified genes or other risk factors (Tang, Stern, Marder, Bell, Gurland, Lantigua et al, 1998). Indeed, the UK incidence estimates note that 6.1% of all people with dementia among BME groups are of younger onset compared with only 2.2% for the population as a whole, which possibly reflects the younger age profile of BME communities living within the UK. However, the 'BME community' is a heterogeneous group composed of different religions, languages and cultures, therefore it is unlikely that genetic factors will account for early onset dementia in all of these ethnic groups. There is also evidence to suggest that measures used to assess cognitive decline, such as the MMSE, in older people from ethnic minority groups have high rates of false positives. There is also the possibility that dementia is under-diagnosed in minority ethnic groups due to their reluctance in accessing services (Eolas, 1999).

According to the 2001 census study there are an estimated 4.6 million (or 7.9 per cent) people in the UK belonging to non-White ethnic groups (ONS; Office of National Statistics, 2003). Indians were the largest of these groups (1.8 %), followed by Pakistanis' (1.3%), those of mixed ethnic backgrounds (1.2%), Black Caribbean (1%), Black Africans (0.8 %), Bangladeshis' (0.5%) and Chinese (0.4%). The remaining minority ethnic groups each accounted for less than 0.5% of the UK population and together accounted for a further 0.9%. The numbers and proportions of older people from BME groups

are rising and will continue to rise for the foreseeable future. However, people from BME groups have a slightly younger age profile. Nine per cent of people from Black and Caribbean ethnic backgrounds are aged 65 and over and six percent of Indians are aged 65 and over. These proportions are expected to expand as people from minority ethnic groups who migrated to the UK in the 1960s and 1970s reach retirement age (ONS, 2003).

Unfortunately there is no data which breaks down the incidence of AD, VaD or any other dementia type illness within BME groups. It is hypothesised that the incidence of VaD would be higher for BME groups since they are a high risk for developing cardiovascular-related difficulties (Chen, 2004; Purandare, Luthra, Swarbrick & Burns, 2007).

1.2.4 Literacy and fluency of English language amongst ethnic minority groups in the UK

One barrier to implementing routine practices for minority ethnic groups is language. There are no consistent or conclusive statistics for rates of literacy and fluency in English within minority ethnic groups. Within migrant groups, older people are least likely to have acquired faculty in the dominant language and integrated into the new nation's culture (Johnson, Owen, Blackburn, Rehman & Nazroo, 2000). In a 1999 survey, among the 50 to 74 age group, 67% of Pakistani women, 46% of Pakistani men, 100% of Bangladeshi women and 64% of Bangladeshi men stated they could speak 'only a little English' (Johnson et al, 2000).

In addition, cognitive testing relies on linguistic competence. Therefore, whether a person understands English or not, tests which are reliant on spoken language to assess cognitive functioning, would limit their usefulness with those who are non-verbal in general. For non-English speakers, clinicians may be dependent on an interpreter;

however, interpreting test instructions is simpler to do than interpreting verbal responses. The reason being that the test responses are the main source of measurement and changing requirements for verbal responses to make them more culturally specific may cause them to be less equivalent to the original test designed for English speakers. Therefore, measures that are not reliant on spoken language can assist the patient to communicate their responses by drawing or pointing to answers (Olsson, 2004).

1.3 Anti cholinesterase Inhibitors (AchI)

The use of AchI for AD has generated controversy in recent years and attracted attention from the scientific community and various non-statutory agencies and consumer groups. In dementia, especially AD, the production of acetylcholine is decreased. As a result, nerve communication becomes less efficient, with consequent problems of cognitive and behavioural functioning. AchI prevent the breakdown of Acetylcholine and can improve nerve function; however the damage to brain cells cannot be halted or reversed.

There are three types of AchIs that are licensed for the symptomatic treatment of people with mild to moderately severe dementia. These are: 1) Aricept (Donepezil), 2) Reminyl (Galantamine) and 3) Exelon (Rivastigmine). In 2003, 77% of prescriptions for AchIs were for Donepezil (NICE, 2007).

1.3.1 Evidence for effectiveness of AchI

Dementia severity is classified in terms of mild (21-26 points), moderate (10-20 points), moderately severe (10-14 points) and severe (0-9 points) based on cut-offs on the MMSE. It is important to be aware of these cut-offs because the evidence base refers to the effectiveness of drug treatments for each of these stages of severity of dementia. However, it is unclear whether RCTs are stratifying their

samples in this manner because NICE has chosen these cut-offs or if NICE has generated their cut-offs on the basis of how samples have been stratified within RCTs. If the latter is the case, it makes logical sense for NICE to only recommend treatments for groups for which they have been proven to be effective. Either way, there is limited information of why these cut-offs have been used as oppose to alternative cut-offs proposed in the literature (Folstein, Folstein & Fanjiang, 2000).

NICE reviewed the clinical and cost effectiveness of AchI for mild to moderately severe AD. The outcomes were explored in terms of cognitive functioning, functional ability, quality of life, behavioural symptoms, adverse events, cost effectiveness, benefits to carers, mortality rates and delay of residential care (NICE, 2007). The use of these drugs is surrounded by controversy and whilst most accept that there is a modest benefit, others insist that they simply do not match up to expectations (Starr & Lonie, 2008; Rodda & Walker, 2009). It was recognised that although measures used to explore the benefits of AchI to carers, functional and behaviour assessments did not show any effects of the treatments. However, this could be due to the lack of appropriate and sensitive measures rather than a lack of effect (NICE, 2007). Furthermore, not all patients do benefit and there are no reliable predictors of who is more or less likely to benefit from these drugs (Starr & Lonie, 2008).

A systematic review to assess the effects of AchIs in delaying the conversion from Mild Cognitive Impairment (MCI) to AD concluded that treatment for periods ranging from less than 4 months to three years, was not associated with any delay in the onset of AD or dementia (Raschetti, Albanese, Vanacore & Maggini, 2007). The eight trials included within this review ranged in the primary and secondary measures used to measure treatment outcomes. Some studies used

global cognitive assessments such as the MMSE as a primary measure and more detailed neuropsychological assessments as secondary measures, such as the Clock Drawing Test, Symbol Digit Modalities Test and Trail Making Test. Although significant treatment effects were not found in the primary efficacy measures, outcomes on secondary measures suggest promising directions for further evaluation of Aricept treatment in patients with MCI (Salloway, Ferris, Kluger, Goldman, Griesing, Kumar et al, 2004). However, Birks & Harvey (2006) report changes of 1 point on the MMSE in some trials, yet this could be accounted for by error variance alone. This suggests that measures such as the MMSE are not sensitive to change and in fact any change of less than 4 points on the MMSE is not informative in terms of the degree of cognitive and functional change.

Birks and Harvey (2006) conducted a meta-analysis to assess whether Aricept improves the wellbeing of patients with dementia due to AD. All unconfounded, double-blind, RCTs in which treatment with Aricept was compared with placebo for patients with mild, moderate or severe dementia due to AD were included. Outcome data covered domains of cognitive function, activities of daily living, behaviour, global clinical state, adverse events and healthcare costs. The MMSE was frequently used to measure severity and was also used as an outcome measure to assess the effect of the treatment on cognitive functioning. People with mild, moderate or severe dementia due to AD treated for a period of 12, 24 or 52 weeks experienced benefits in cognitive function, activities of daily living, and behaviour (Birks & Harvey, 2006).

A meta-analysis for treatment of VaD with AchI concluded that they resulted in small benefits in cognition, with uncertain clinical significance in patients with mild to moderate VaD (Kavirajan & Schneider, 2007). However, the study highlighted the challenges in

designing clinical trials for vascular dementia. The heterogeneity of the clinical group leads to differing clinical presentations and course of disease, therefore they may respond differently to the drugs. Furthermore, the small effects shown in trials for vascular dementia could be accounted for by the effects on coexisting AD, rather than the VaD (Kavirajan & Schneider, 2007). In addition, the outcome measures used are ones that are designed for AD populations, therefore may not be sufficiently sensitive to clinical changes in VaD patients.

1.3.2 Minority ethnic groups excluded from clinical trials

RCTs are the most effective method for evaluating healthcare treatments (Altman, 1996). However, there are concerns about the generalisability of findings to ethnic minority groups (Hussain-Gambles, Atkin & Leese, 2004). A review of exclusion criteria used in RCTs found that many had blanket exclusions for ethnic minority groups without any justification (Britton, 1999), even in the UK (Heiat, Gross & Krumholz, 2002). This has important repercussions regarding the safety and efficacy of new drug use in these groups, especially since drug metabolism, concurrent diseases and counter-indications have been shown to vary considerably between different ethnic groups (Krecic-Shepard, Park, Barnas, Slimko, Kerwin & Schwartz, 2000).

Apart from being poor science this undermines the UK Government's plan for addressing inequalities, and its core principle of providing culturally appropriate and accessible care for all in the NHS (Department of Health, 2000). Hussain et al (2004) suggest that the exclusion of ethnic minority groups who do not speak English amounts to institutional racism and efforts should be made to make research-related information and measures accessible to these groups also (Hussain-Gambles, Atkin & leese, 2004). However, there

are challenges to including people who do not speak English in AchI trials. Although solutions are possible, they are complicated. The lack of cross-culturally appropriate cognitive measures for dementia is a significant barrier. Clinical trials of AchIs use quantitative cognitive assessments, which have only been validated within western cultures, as primary and secondary outcome measures. It is assumed that comprehension and expression of English would be required to administer the assessments; therefore those who are not fluent in English would need to be excluded.

Including people who are non-fluent in English may be unethical, especially if outcome measures are administered with the knowledge that they will not lead to any results which would be scientifically or clinically valid due to increased error variance (Emanual, Wendler & Grady, 2000). This is especially concerning in the context of drug trials, which can result in adverse affects. If non-English speakers are to be included within clinical trials then more appropriate cognitive measures need to be developed, that are cross-cultural and are equivalent to the measures used for English speakers from the majority ethnic groups. This could aid in making more accurate comparisons of dementia disorders across different countries and cultures (Gibbons, van Belle, Yang, Gill, Brayne, Huppert et al, 2002).

1.4 National Institute for Health and Clinical Excellence (NICE)

NICE decides what treatments are supplied on the NHS in England and Wales, and are asked to look at particular drugs and devices when there is confusion or uncertainty over the value of a drug or device, or when prescribing practices vary across the country. NICE have been under pressure to approve AchI for AD since they were first introduced in 1996 in the US. Between 1997 and 2000 various inconsistencies became apparent in the prescription of Aricept. Patients could receive different treatments depending on where they

happen to live, and the clinician treating them rather than on the state of their health (Doyle, 2001). The Merton, Sutton and Wandsworth Health Authority was cited as an example by Doyle (2001). In 1997 the authority concluded that the benefits of AchI were insufficient to warrant funding. However, some general practitioners were prescribing the drug in primary care and patients were buying it privately (Doyle, 2001). There is evidence that due to abundance of guidance from various institutes, clinicians may not view them as being worthwhile (llife, & Manthorpe, 2007).

In general NICE is widely admired as a model for decision-making on the introduction of new technologies to a healthcare system because they balance clinical effectiveness with cost effectiveness, which is essential to ensure equity in service provisions (Syrett, 2007). Cost-effectiveness analysis is a tool for arriving at a "cut-off" point for what is valuable, and what society should pay for, and what should not be covered. The Quality Adjusted Life Years (QALY) is the model used by NICE to compare different drugs and measure their cost effectiveness relative to clinical effectiveness.

1.4.1 NICE guidance for treatment of AD with AchI

AchI had first been the subject of NICE Guidance in 2001. These had approved AchI for mild and moderately (a score over 12 points on MMSE) severe AD only. NICE updated their guidance and provided 'restricted' guidance for clinicians prescribing AchIs for AD. NICE based this decision on evidence suggesting that AchIs were not a clinically effective or cost effective treatment for people with mild or severe AD and that evidence supports effectiveness only with people with moderate AD (NICE, 2007). Updated guidance stated that patients were eligible for treatment with AchI only if their cognitive functioning falls within the 'moderately' severe range (10-20 points) on the MMSE (NICE, 2007). Therefore patients within the 'severe' (0-

9 points) and 'mild' range (21-26 points), as measured by the MMSE, should not be prescribed the medication even if clinicians disagree with the guidance (NICE, 2007). The MMSE is also used to monitor patients on this medication, and the medication is to be stopped if it does not appear to be having an effect, or if the patient's scores fall below 10 points on the MMSE.

It was inevitable that restricting access to expensive drugs that have modest benefits would lead to public protest (Doyle, 2001; Godley, 2006). This is even more likely to happen in an area where there is already a lack of alternative preventative drug treatment options available (Melzer, 1998). Thompson et al (1997) argued that the recommendations were wrong because the AchIs are all that doctors have to offer. Subsequently in 2006 NICE was faced with its first High Court challenge (Dyer, 2006; Kmietowicz, 2006; Syrett, 2007). A pharmaceutical company (Eisai) which is the licence holder for Aricept applied for a judicial review to challenge NICE's decision. The judge ruled that NICE were not irrational in concluding that there is no cumulative benefit to patients after six months treatment, that they did take into account benefits the drugs bring to carers, and they did not breach principles of procedural fairness.

NICE cannot be held solely accountable for insufficient evidence that the drugs are cost effective in early stages of dementia (Godley, 2006). Furthermore, drug companies and their marketing strategies can open the NHS to exploitation through their efforts to gain unrestricted access to the NHS as part of a free market. NICE is in a position to regulate this using evidence (Melzer, 1998; Kmietowicz, 2006). Drug companies have been known to misinform the public about the efficacy of AchI (Melzer, 1998; Doyle, 2001). It is possible that this initial misinformation contributed to the impetus to challenge NICE on their decision and accuse them of ageism and stigmatising people with dementia (O'Brien & Ballard, 2001; Dening, 1992). Other

key contributory factor are misconceptions of NICE's role and an over-estimation of the effectiveness of the drugs for any stage of dementia (Doyle, 2001).

1.4.2 NICE guidance & non-English speakers

NICE was found to have acted unlawfully in one respect. Both guidance issued by the institute and the decision of the appeals panel were found to be flawed in that there had been a failure to show proper consideration to NICE's statutory duties under the Race Relations Act (1976) and the Disability Discrimination Act (1995), to promote equal opportunities and to eliminate discrimination. This difficulty arose as a result of the failure of the guidance to address the position of those for whom English is not a first language, those with learning disabilities, and those with sensory difficulties such as individuals who are blind, deaf or non-verbal. The administration and interpretation of the MMSE is problematic within these populations. What is also of concern was NICE's decision to advocate that the MMSE should be used as a measure to decide about treatment eligibility. Although NICE advises that healthcare professionals should not rely on the MMSE score alone in these circumstances, Cerejeira & Mukaetova-Ladinska (2007) argue that one should not rely on the MMSE score alone in any circumstance.

In acknowledgement of this, NICE stated that these groups should have equal access to treatment and should not be discriminated against if unable to complete the MMSE (NICE, 2007). The guidelines suggest that in these instances alternative measures should be selected, although, they do not provide suggestions for alternative assessments. This is probably because there are no validated measures which are appropriate for non-English speakers, which are also equivalent to the MMSE, and have proven utility in decision making regarding treatment with AchI.

Following the judicial review, NICE was directed to amend their guidance. Their amended guidance stated that if the patient is judged to have AD of moderate severity, as determined by other forms of assessment, they should be prescribed AchI. However, this merely reinforced the limitations of the guidance as there are no objective means of interpreting what AD of moderate severity would look like in terms of cognitive, behavioural, functional or neurological characteristics. The MMSE is what defines the stages of severity and research which aims to illuminate what the parameters of the 'moderately severe' range is in terms of patient characteristics and cognitive profile is of critical importance if these guidance are to remain applicable in routine clinical practice, in general, and for people who do not speak English.

1.5 Mini-Mental State Examination

Cognitive screening is an integral part of assessing for cognitive decline to establish the health care needs of older adults (Wilkins, Wilkins, Meisel, Depke, Williams & Edwards, 2007). The MMSE (Folstein, Folstein & McHugh, 1975) is the most frequently used and highly cited assessment for screening cognitive functioning in older people (Tombaugh & McIntyre, 1992; MacKenzie, Copp, Shaw & Goodwin, 1996; Davey & Jamieson, 2004; Nilsson, 2007; Mitchell, 2009). Although it was originally developed in hospital settings, it has been applied widely in clinical practice, clinical research and epidemiological studies (Huppert, Cabelli & Mathews, 2005). It was devised as a short, convenient, quantitative assessment of cognitive performance and was viewed as being a more suitable measure for older people in comparison to alternative more lengthy assessments due to their difficulties in sustained attention. There is a preference for brief, sensitive screening measures which are easily accessible, inexpensive and which can be administered by non-specialists. Thus, making them more practical for routine assessment in primary care

settings, although many in these settings still consider the MMSE too time consuming to administer (Glasser, 1993; Mitchell, 2009).

The distribution of scores across the eleven items of the tests are not weighted equally, therefore the measure does not screen for decline equally across these cognitive domains. Thus, it is not sensitive to impairments in memory, which is the most affected of cognitive domains in AD. A criticism of this measure is also that it does not directly measure executive functioning, which is another key domain affected within the dementia population. Given that it is a screening measure, it can be forgiven for not being a comprehensive assessment of any domain; however this is not always appreciated in practice. Tombaugh and McIntyre (1992) in their review concluded that in general the MMSE fulfilled its original goal of providing a brief screening tool that assesses the severity of cognitive decline and documents change over time. However, content analyses revealed the MMSE was highly verbal, and not all items were equally sensitive to cognitive impairment. Items measuring language were judged to be relatively easy and lacked utility for identifying mild language deficits. Overall, MMSE scores were affected by age, education, and cultural background, but not gender (Tombaugh & McIntyre, 1992; O'Connor, Pollitt, Hyde, Fellows, Miller, Brook et al, 1989; Brayne & Calloway, 1990; Fratiglioni, Jorm, Grut, Viitanen, Holmen, Ahlbom et al, 1993; Wind, Schellevis, Stavernen, Scholten, Jonker & Van Eijk, 1997).

Nevertheless, a score between 0-9 indicates severe cognitive decline, 10-20 moderate, 21- 26 mild and 27-30 is considered to be within the 'normal' range (NICE, 2007). These distinctions between separate levels of severity have been found to be useful for clinical and research purposes. Staging in this way has enabled clinicians to categorise and hence communicate information rapidly about the

disease. Staging is also required when deciding about treatments that have been approved for different levels of disease severity (NICE, 2007). In research, staging has been used to select 'homogenous' or comparable patient samples.

However, there are two fundamental theoretic and empirical misconceptions underlying current clinical and research practices. Firstly, the MMSE is a screening tool used to identify cognitive difficulties, therefore by definition it identifies symptoms of cognitive decline to support the need for further neurological, neuropsychological, functional and laboratory examinations (Folstein, Folstein & McHugh, 1975). It should not be used on its own to make a diagnosis of any neurological pathology or specify the underlying clinical condition which is causing cognitive decline (Field, Jackson, Hassett & Pattison, 1995; Bird, Canino, Stipec & Shrout, 1987; Mitchell, 2009). In addition, it is not sensitive enough to identify milder levels of cognitive decline in dementia (Peterson, Stevens, Ganguli, Tangalos, Cummings & DeKoski, 2001; Mitchell, 2009).

In the original paper, Folstein and colleagues highlighted that a key advantage of the MMSE was that the items measured had clearer implications on patient's capacity to care for themselves and manage daily affairs. Whether or not the MMSE scores always correlate with functional abilities is debateable, but the measure was viewed as being useful for identifying social support that the patients may need, not advocated as a quantitative measure whose cut-offs could be used to decide on eligibility for drug treatments. It could be argued that the utility of the MMSE has advanced since the measure was first developed, but the evidence has repeatedly called for more appropriate use and in particular for age and education adjusted norms. This has not made an impact on clinical guidance. NICE in their guidance do state that age and education should be taken into

account when interpreting scores. However, without normative data it is unclear how clinicians should interpret scores.

Furthermore, although a cut-off score of 24/30 is recommended by Folstein and colleagues, NICE guidance use 26/30 as a cut off. Folstein and colleagues (1975) state that psychiatric in-patients with dementia invariably scored 23 points or less out of 30, compared to normal controls, all of whom scored above this cut-off point. However, in-patient populations are likely to be experiencing difficulties that are in the more severe end of the spectrum compared to those who remain at home (O'Conner, Pollitt, Hyde, Fellows, Miller, Brook & Reiss, 1989). Therefore, these cut-offs are likely to discriminate less accurately between those with dementia in normal community populations. Numerous other cut-offs have been calculated from Receiver Operating Characteristic (ROC) analysis of specific populations together with adjustments for age and education (Kahle-Wrobleski, Corrada, Li & Kawas, 2007; Mitchell, 2009). This was in response to the ample evidence that age and education were confounding factors (Crum, Anthony, Bassett et al, 1994; Grigoletto, Zappala, Anderson & Lebowitz, 1999; Mitchell, 2009).

There are additional sources of psychometric instability for the MMSE. Classical Test Theory indicates that obtained scores consist of two components: the true score and the error variance. Longitudinal testing is a hallmark of dementia assessment and the MMSE is administered repeatedly (Pelosi, McNulty & Jackson, 2006; NICE, 2007), this can result in practice effects (Hinton & Withers, 1971). The limited use of standardised assessment instructions (Meyers & Meyers, 1995) results in differences in administration practices across clinicians. This results in reduced inter-rater reliability, even amongst specialists (Davey & Jamieson, 2004).

There are also item-specific problems which can introduce error variance. For instance the attention subtest has two very different items and the patient is expected to complete only one of them. If they will not or cannot perform the calculations item, they have the option of spelling 'world' backwards (Davey & Jamieson, 2004). The item administered can vary from different clinicians and also there is evidence to suggest that the backward spelling task is easier, and so less sensitive in comparison to calculations (Galasko, Klauber, Hofstetter, Salmon, Lasker & Thai, 1990). Therefore, a change in scores could reflect psychometric instability, which has been found to be the case in a normative sample (Olin & Zelinski, 1991). Furthermore, a change in the environment in which the test is conducted may affect responses to the orientation subtest. A client who is asked to name the geographical location of where they are or name the building they are in, may respond incorrectly when assessed during an outpatient appointment compared to when being assessed in a familiar environment (i.e. their own home).

Thus, although the cut-off scores can provide useful screening information to inform clinical decision making for further assessment, there is no consistent findings within the research literature to evidence that it is scientifically reliably, valid or clinically meaningful to use these cut-offs alone to decide on treatment eligibility. Unfortunately, the psychometric limitations of the MMSE have not diminished its popularity and paradoxically, research has attempted to extend its usage beyond that of a dementia screening (Olin & Zelinsky, 1991).

Secondly, although theoretically it would seem sensible that all those who fall within the mild, moderate or severe range would present with comparable cognitive and functional difficulties to their counterparts, this is not always observed in clinical practice. The

'mild', 'moderate' or 'severe' ranges lack construct validity and criterion validity. Patients can vary on what items they score lower points on and still fall within the same level of severity. Therefore, research, especially RCTs, that categorises participants when trying to evaluate the efficacy of drugs for specific levels of severity are confusing. When selecting samples, RCTs should include functional, behavioural, alternative cognitive and neurological measures to inform their inclusion criteria and ensure that the sample is comparable in terms of level of severity, rather than relying on the MMSE alone.

The use of the MMSE as an outcome measure also demonstrates a lack of appreciation for the multiple sources of error variance. The many confounding factors influencing the MMSE cause it to be an unstable measure and consequently makes it difficult to know if an increase in scores were indicative of a genuine treatment effect, or if a reduction in scores was indicative of decline. Therefore, outcomes need to be balanced by a design which attempts to control adequately for confounding variables. This is especially important if drug treatments will then be approved only for specific levels of severity (NICE, 2007). A failure to do this could lead to erroneous conclusions being made both in research and clinical settings (Olin & Zelinski, 1991). However, this is by no means an easy task, which is perhaps why the dominant trend has been to use the MMSE. Patients with dementia of any type are not homogenous and without a better understanding of the underlying aetiology of the disease, to inform diagnostic criteria, it will be a challenge to identify a measure that will fit for all. Furthermore, cognitive, behavioural and functional abilities are strongly influenced by cultural, educational and life experiences and it is not possible to control for all of these variables, especially within a much older sample.

1.6 MMSE and BME groups

Research suggests that where the original MMSE has been used with ethnic minority groups there has been an increased risk of false positives (Parker & Philip, 2006), hence substantially increasing the risk of them being diagnosed with dementia. In one study 6% of non-impaired White people and 42% of non-impaired Black people were wrongly classified by the MMSE (Fillenbaum, Heyman, Williams, Prosnitz & Burchett 1990). One factor which could account for this variance is years of education. There is a well-documented influence of education on performance on cognitive tests (Tombaugh & McIntyre, 1992). However, even when years of education have been accounted for, there remains an effect, related to ethnicity and culture. To support this further, people from the White ethnic group in USA and UK performed differently on MMSE, which demonstrates a clear effect of culture rather than race (Gibbons et al, 2002).

One approach to developing equivalent measures to the MMSE has been to adapt items for cultural equivalence. Helms (1997) suggested that to control for cultural bias in neuropsychological assessment the tests should be matched according to (1) functional equivalence (extent to which the tests measure the same psychological constructs with equal accuracy within these groups), (2) conceptual equivalence (level of familiarity with the test items); (3) linguistic equivalence, (extent to which language used in the tests has equivalent meaning across cultures); (4) psychometric equivalence (extent to which the tests measure the same thing at the same level across cultures); (5) testing condition equivalence; (6) contextual equivalence, (evidence that the cognitive ability being assessed is comparable across environments; and (7) sampling equivalence (samples of subjects representing cultural groups should be comparable).

In an attempt to develop culturally appropriate assessments the MMSE has been translated in various languages, such as Spanish

(Escober, Burnam, Karno, Forsythe, Landsverk & Golding, 1986), Chinese (Katzman, Zhang, Qu, Wang, Liu, Yu et al, 1988) Korean (Park, Park & Ko, 1991), Hindi (Ganguli, Ratcliff, Chandra, Sharma, Gilby, Pandav et al, 1995), Japanese (Larson et al 1992), Gujarati (Lindesay, Jagger, Mlynik-Szmid, Sinorwala, Peet & Moledine, 1997a), Urdu (Rait et al, 2000) and Sinhalese (de Silva & Gunatilake, 2002). Many of the studies involve translation of the instructions and test items, and some aspects of the tests are modified to be more culturally appropriate. For example in the Sinhalese translated MMSE for the Sri Lanka population, they replaced 'season' with 'time of day' for orientation to time, because there are no major seasonal variations in Sri Lanka (de Silva & Gunatilake, 2002). Translations require equivalence in terms of word meaning and therefore in order to do this accurately measures are translated and then back translated to ensure that correct translations are being used. However there are occasions when another language does not have an equivalent word and therefore it will be replaced with an alternative word. Hence, reducing the equivalence of the MMSE with translated measures.

The studies that are based in the UK (Lindesay et al, 1997; Rait et al, 2000), have small sample sizes, therefore the measures are not validated in the UK and results cannot be generalised, hence the lack of normative data. Furthermore, clinicians who are English speaking would not be able to administer translated MMSEs and would still require access to an interpreter. It is also difficult for clinicians to supervise translators, ensuring that correct translations and procedures are being followed (Nell, 1999). One would have to question the appropriateness of interpreters administering cognitive assessments with the absence of specialist training, knowledge and the skills required to assess the outcomes. A trained clinician is able to consider all possible sources of score variance such as, vision,

mood, environment, demand characteristics, fine motor skills, speech and language difficulties to name a few.

In addition, it is unclear whether translated MMSEs are actually utilised within services once they have been developed. There is also a continuing debate over the suitability of instruments developed in one culture (e.g. MMSE) being used in another culture. Factors such as language, literacy and ethnicity, acculturation, age and education have been shown to have an effect (Tombaugh & McIntyre, 1992;Baird, Ford & Podell, 2007; Boone, Victor, Wen, Razani & Ponton, 2007).

Nevertheless, there remains a need to develop standardised assessment procedures that are objective, especially when making decisions about treatment and to monitor change following drug treatments. Despite the limitations of using existing cognitive tests, it may be easier to adapt existing tests, especially when they are needed imminently in clinical settings such as when deciding if a person is eligible for medication. In addition, translated MMSEs can be used in combination with cognitive tests that do not require spoken language and their combined evidence could provide more reliable insights into a patients cognitive functioning. The combination of multiple assessment modalities is best practice in neuropsychological assessment (Lezak et al, 2004).

1.7 Cross Cultural Neuropsychology

Chesters (2007) suggests that only a limited kind of neuropsychology, appropriate to only a fraction of the world's population, is being presented to the rest of the world as if there could be no other kind of neuropsychology. Much can be learned from studying diverse cultural groups, to add to our understanding of dementia, rather than depending only on research that tends to be Eurocentric. Cross-cultural analysis may give us a better

understanding of brain organisation and treatment efficacy, and we are limited because our knowledge is based on research evidence from western civilisations only (Ardila, 1993). What is also urgently needed is the increase in the knowledge base of, and provision of appropriate tools for clinicians to enable them to carry out culturally competent and clinically relevant neuropsychological evaluations (Perez-Arce, 1999). Indeed, if the field is to remain scientifically credible, it needs to address the effects of language and culture on cognition, and provide access to neuropsychological assessment and rehabilitation services to all (Ivnik, 2005).

In the last decade there has been an increase in the number of publications addressing the relevance of cultural factors to performance on neuropsychological measures (Agranovich & Puente, 2007). Factors such as degree of acculturation, cultural experience, literacy level and quality of education influence test performance (Manley, Jakobs, Sano, Merchant, Small & Stern, 1998). There are also different norms in different cultures in terms of attitudes towards testing (Agranovich & Puente, 2007). It is also suggested that cultural difference can also affect lateralisation of language and spatial disturbance (Ardila, 1995) and thus have a profound effect on non-verbal behaviour and language.

The increase in demands for neuropsychological assessment of persons with limited or no English language background has been the impetus for developing instruments written in the patient's language and standardised for persons from a specific cultural or for a specific language (Uzzell, Ponton & Ardila, 2007). The use of interpreters is only a second best partial solution (Ardila, 1995; Lezak et al, 2004). The next important goal for neuropsychological assessment should be the dissemination of evidence-based language and culture appropriate neuropsychological examination techniques and skills

(Lezak et al, 2004). There are large variations between individuals and diseases and it is unlikely that a single test will be appropriate for everyone; however examination in specific areas should lead to some simplification and generalisation in procedures. Therefore, a series of relatively short fixed batteries designed for particular disorders, specific domains (e.g., memory, attention, executive functions), that answer specific questions and meet particular needs are what we should be working towards (Lezak et al, 2004).

1.8 Aims

Preliminary data was collected from participants whose first language was English to establish the utility of the cognitive measures, to aid in decision making regarding eligibility for treatment with AchI. Therefore the data collected will not, and did not intend to, provide information regarding the cultural equivalence of the chosen measures. It is hoped that the conclusions drawn from this study would be used to inform the design of a study in the future to explore the cross-cultural utility of these cognitive tests in establishing treatment eligibility in a non-English speaking sample.

The hypotheses were:

1. There would be a significant correlation between participants' total scores on the MMSE and their total scores on the Rey Complex Figure Test (RCFT), Raven's Colour Progressive Matrices (RCPM), Symbol Digit Modalities Test (SDMT), Brixton test, the Clock Drawing Test (CDT), and the Colour Trails Test (CTT).

The participants' performance on the RCFT, RCPM, SDMT, Brixton test, CDT, and the CTT would differentiate those eligible for treatment with AchI (10-20 score on MMSE) from those who are not.

2. Extended Methodology

2.1 Ethical Considerations

An application for ethical consideration was made to the National Health Service (NHS) Central Office for Research Ethics Committees. Ethical approval was obtained prior to commencement of the study. Following this, further approval was obtained from the Research and Development departments of the two NHS sites. Three key ethical considerations were highlighted by the ethics committee, 1) to ensure that participants had the capacity to consent and this was assessed formally by the lead researcher, 2) to maintain participant anonymity and 3) to allow participants to take breaks as and when needed from testing. The ethical issues raised were taken into consideration in the methodology of this study (see Appendix 2 for NRES approval).

2.2 Sample size and power

A priori sample size calculation was conducted for this study. The sample size is calculated on the basis of the desired effect or power that is hoped to be achieved. A power of .80 is the standard adequacy level (Dancey & Reidy, 2007). The effect size is to be informed by previous research in the area (Field, 2000). However, this can be an arbitrary process if there are no previous studies which have explored the strength of a relationship between variables of interest. In addition, when a study involves the exploration of the strength of a relationship between multiple variables or measures, the expected effect can vary within a study, for each measure. For this study, there were no prior studies which had explored the correlation between the MMSE and the RCFT subtests, RCPM, SDMT, Brixton test or the CTT. However, previous research has reported that the CDT correlated moderately/ highly (.41 to .80) with the MMSE (Strauss, Sherman & Spreen, 2006). Therefore, it would be appropriate to conduct a priori sample size calculation based on a large expected effect. However, even though the MMSE has not been

correlated with the Brixton test, the MMSE is reported to be a poor measure of measure of executive functioning, therefore a small relationship would be expected between the MMSE and the Brixton test. Nonetheless, due to the nature of this study, which is interested in identifying measures that can aid in decision making about treatment with AchI, a strong effect is desired. Therefore, for a large effect (0.5, Cohen, 1992), standard adequate power of .80, and significance level of $P<0.05$, the minimum sample size required was 29 participants. This was calculated using G-Power analysis (Erdfelder, Faul & Buchner, 1996).

2.3 Exclusion criteria

Patients were excluded if:

1. They were blind or deaf or had difficulties with fine motor skills: This is because the assessments could not be administered to these patients in a standardised way. The tests that were administered required participants to be able to draw and write therefore they needed to have intact fine motor skills. Participants who were blind or deaf were also excluded as these modalities were needed to carry out the tests. This was assessed by asking the participant prior to commencing the study. If they indicated that they experienced difficulties in these areas they were excluded.
2. They did not consent to take part.
3. They did not have the capacity to consent
4. Anxiety or mood related difficulties: Mood related difficulties such as depression and anxiety can impact on test performance. Therefore, if participants indicated that they felt distressed or were unable to concentrate due to mood related difficulties they were excluded from this study. No formal psychiatric assessment was conducted with participants. In addition, participants who appeared distressed during testing would be excluded, especially

if this was likely to have impacted on their performance and would also have been unethical to continue testing. This was assessed using observation.

5. Participants who had a recent (1 year) head injury were also excluded following the initial visit. This was assessed on the basis of participants self-report.

2.4 Measures

2.4.1 The University of San Diego Brief Assessment of Capacity to Consent (UBACC, Jeste, Palmer, Appelbaum, Golshan, Glorioso, Dunn et al, 2007)

One reason for the paucity of research in dementia is ethical concerns around capacity to consent (Agarwal, Ferran, Ost & Wilson, 1996; Warner & Nomani, 2008). The assessment of capacity has been codified in law (DOH, 2005) and there is increasing demand for formal capacity assessments of potential participants with dementia (Kim & Caine, 2002). What is also concerning is that many individuals with dementia participate in research but are likely to lack capacity (Warner, McCarney, Griffin, Hill & Fisher, 2008).

Participants should be able to understand salient information provided to them, retain this information for sufficient time to be able to weigh the information in the balance, and to form a decision (without coercion), and to communicate that decision. However, there is no 'gold standard' capacity assessment and there is considerable heterogeneity within diagnostic groups. Factors (such as cognitive impairment) that have the most significant association with impaired capacity explain no more than 25% of the variance (Jeste & Saks, 2006; Church & Watts, 2007). The MacArthur Competency Assessment Tool for Clinical Research [MacCAT-CR] (Appelbaum & Grisso, 2001) and the University of California, San Diego Brief Assessment of Capacity to Consent [UBACC] (Jeste et al, 2007) are

two frequently used instruments used to assess for capacity to consent in research settings. Kim and Caine (2002) used ROC analysis to examine the utility of the MMSE in discriminating capacity status, against the MacCAT-CR. They found that the MMSE did not discriminate capacity status well, which is supported by a more recent study (Warner et al, 2008). They suggested that a cut-off of 26 would have sensitivity of 91-100% and low specificity of 35.7%. Indeed, this leads to suggestions that people with moderate and severe dementia as measured by the MMSE, have limited, if any capacity to give informed consent (Sugarman, Roter, Cain, Wallace, Schmechel & Welsh-Bohmer, 2007).

Carpenter et al (2000) in their study found that participants should not automatically be excluded from studies if they perform poorly on the MacCAT-CR, as participants are able to provide informed consent if additional educational information is provided (Carpenter, Gold, Lahti, Queern, Conley, Bartko et al, 2000). If people were excluded purely on the basis of this measure, which does not take into account the unequal thresholds of competence across different studies with varying risk-benefit ratios, dementia research would be halted (Warner & Nomani, 2008). Carpenter et al (2000) suggested that retention of key elements, rather than detailed knowledge is critical during research. The key elements include awareness that participation is voluntary, and that withdrawal can be done without penalty. Remembering the name of drugs and the full list of adverse effects is less important.

When appraising the literature for guidance regarding seeking informed consent for this study, it was apparent that the focus of the literature base was on participation in RCTs. Therefore, it is difficult to apply this evidence to studies which do not involve drug treatments. Nonetheless, a formal assessment of capacity was incorporated in

this study and the measure chosen was the UBACC (Jeste et al, 2007). The reason being that it takes less time to complete compared to the MacCAT-CR, and has standardised procedures for administration and scoring. The items or questions can be re-worded to make them more applicable to the specific study. However, the items reflect the key areas that need to be assessed (see Appendix 7). A response is scored either 0 (in-appropriate response), 1 (warrants further information to assist participant, or further assessment of capacity is required), or 2 (appropriate response indicating capacity to consent). An additional advantage of this measure was that the participant is allowed to refer to written information (such as the participant information sheet (see Appendix 6) to answer the questions, hence they do not need to rely on memory alone. This is more appropriate for participants with dementia.

2.4.2 Demographic questionnaire

This information was collected from carers only when the participant provided consent. This information was documented separately for each participant on a demographic questionnaire sheet (see Appendix 10).

2.4.3 **MMSE** (Folstein et al, 1975):

The MMSE measured orientation to time and place (date, time, town – 10 points), memory (remembering three objects immediately – 3 points; a recall trial– 3 points), comprehension (following instructions – 3 points), language (naming objects -2 points; following commands, reading and writing a sentence – 3 points), attention (subtract 7 from 100 or spell WORLD backwards – 5 points) and visual spatial abilities (copying a drawing- 1 point). The average time taken to complete the test was 15 minutes and the scores range from 0 to 30 points (0-9 severe, 10-20 moderate, 21-26 mild, 27- 30

normal range). A lower score indicated cognitive impairment (Folstein et al, 2000). Many of the items or domains assessed in the MMSE are similar to those assessed in neuropsychological tests. However, the MMSE is considerable to be less sensitive. As a screening measure, it has good criterion validity for people with severe to moderate levels of cognitive decline (100% sensitivity and 85% specificity), but is less predictive of people with milder levels of cognitive decline. A cut-off of 23 and 24 points optimizes its predictive validity (Tombaugh & McIntyre, 1992).

The internal consistency of the MMSE tended to be lower for community samples compared to clinical samples, ranging from a coefficient of .77 (Holzer, Tischler, Leaf & Myers, 1994), to a modest coefficient of .62 (Tombaough, McDowell, Kristjansson & Hubley, 1996). This is likely to be because the MMSE is not a sensitive measure of cognitive decline in community samples, as they score higher on the MMSE.

The MMSE is also designed to assess a variety of cognitive domains, therefore resulting in lower alpha levels in comparison to more detailed neuropsychological assessments (Strauss et al, 2006). Concordance rates between the individual MMSE tasks (orientation, memory, attention) and neuropsychological measures of the corresponding domains can be low (Jefferson, Consentino, Ball, Bogdanoff, Kaplan & Libon, 2002).

Age and education affect performance on the MMSE. Lower educational levels increase the likelihood of misclassifying non-impaired people as cognitively impaired. Furthermore, race, ethnicity and language also affect MMSE performance. Despite this there are over 100 translations of the MMSE, but few have been validated and lack normative data (Auer, Hampel, Moller, & Reisberg, 2000). In addition there is considerable variability in reliability estimates, which

suggests that for conditions which can fluctuate or there are high rates of comorbidity, the MMSE is less reliable (Jorm, Scott, Cullen, & MacKinnon, 1991). For heterogeneous clinical populations, such as those with dementia, the test-retest reliability ranges from .56 to .98 (Folstein, et al, 2000). Studies have reported adequate levels of inter-rater reliability in clinical and community samples, with reliability coefficients ranging from .83 to .95 (Molloy, Alemayehu, & Roberts, 1991). However, inter-rater reliability estimates in research settings are likely to be higher than those found in clinical practice, due to variations in administration practices. Therefore, some caution is needed when interpreting MMSE scores in clinical practice.

2.4.4 Rey Complex Figure Test (RCFT, Meyers & Meyers, 1995):
The RCFT consists of four subtests assessing five mental processes, 1) visual spatial construction ability; 2) speed of processing of visual information, 3) immediate memory, 4) delayed memory, and 5) recognition memory. All standardised instructions were communicated verbally and although the participant could speak during administration, the test did not require spoken language by the participants to be completed. The immediate memory trial is administered after three minutes. Then, 30 minutes after the participant has completed the copy trial, the delayed memory trial begins. The recognition trial commenced immediately after the delayed memory trial was completed. This involved them looking through 24 geometric figures, 12 of which were correct elements of the figure they were asked to copy and 12 were distracters. The participants were asked to indicate which of the 24 items they recognised as being from the stimulus figure (see Appendix 10 for standardised instructions given to each participant).

The standardised approach to scoring within the manual was used. Here the complex figure is divided into 18 units and each unit is

scored separately for both accuracy and placement. A score of 0, 0.5, 1 or 2 was assigned to each unit of the figure based on accuracy and placement criteria. Thus, raw scores range from 0 to 36. For the speed of processing trial, participants were timed on how long they took to copy the drawing. Higher time to copy scores indicated reduced speed of cognitive processing. Lower copy scores indicated reduced visual-perceptual and visuomotor integration skills. Reduced immediate and delayed recall scores suggest reduced visuospatial recall ability, while recognition total correct measures the ability to retrieve visuospatial material when given retrieval cues (Meyers & Meyers, 1995). There are no cut-off scores and participants scores were compared to norms based on their age (Meyers & Meyers, 1995).

Research analyzing inter-rater reliability of scoring of drawings, indicate that the manual instructions are sufficiently detailed to obtain good inter-rater reliability when the raters are experienced clinicians (Strauss et al, 2006; Allen, Brechin, Skilbeck & Fox, 2007).

Evaluation of internal consistency demonstrated adequate reliability for the copy condition (.60) and higher for the recall conditions (.80), which suggests that all the details of the figure tap into a single factor (Fasteneau, Bennett, & Denburg, 1996). The test is sensitive to individuals with disorders which are known to affect memory and executive functioning such as dementia type illnesses (Ardila, Lopera, Rosselli, Moreno, Madrigal, Arango-Lasprilla et al, 2000; Strauss et al, 2006). Correlations with tests of executive functioning were moderate, providing evidence of convergent validity, and correlations with tests of general cognitive ability were low, providing evidence for discriminant validity (Strauss et al, 2006). However, age influences copy performance, especially after 70 years (Tombaugh & McIntyre, 1992; Boone, Lesser, Hill-Gutierrez, Berman & D'Elia, 1993), and is

also positively correlated with education. However, Chervinsky et al, (1992) found that correlations with IQ are modest, which suggests that the RCFT provides a large amount of information not accounted for by education and IQ (Chervinsky, Mitrushina, & Satz, 1992).

2.4.5 Raven's Colour Progressive Matrices (RCPM, Raven, & Court, 2003):

This measure assessed inductive reasoning ability in the visual modality. The test consisted of 36 items, grouped into three sets (A, Ab, B) of 12 items each. In each item participants were presented with an incomplete design and six alternatives among which one must be chosen that best completes the design. Every correctly solved item resulted in a score of one. Items became progressively more difficult and easier items at the beginning of each set acted as learning experiences to be used to inform performance on more difficult items as the test progressed. This indirectly assessed intellectual efficiency by observing if learning increased efficiency and accuracy (Mills, Ablard, & Brody, 1993). The standardised instructions within the published manual were used to administer the tests (see Appendix 10). The test did not require spoken language from participants and verbal instruction from the administrator was kept to a minimum (Zaidel, Zaidel, & Sperry, 1981). Scores ranged from 0 to 36 points, and a lower score indicated greater impairment. This was an untimed test and took up to 25 minutes to complete.

The RCPM is moderately correlated with other tests of intelligence, such as the Wechsler, Stanford-Binet tests, and the NART, which suggests that it has concurrent validity (Strauss et al, 2006). However, age is correlated with the Ravens test, and increasing age is related to decline in performance. A number of studies have also demonstrated strong correlations with working memory with older people, which suggest that memory loss, can be a confounding factor (Salthouse, 1993). Speed of processing also correlates with test

performance, with higher scores associated with rapid processing (Bates & Rock, 2004). However, people with dementia experience global decline in their cognitive functioning, therefore, it is difficult to identify a test which will not be confounded by the multiple domains that are affected. For instance, difficulties with executive functioning can impair the participants' ability to comprehend test instructions. In clinical settings, when a comprehensive neuropsychological assessment is completed with a dementia patient, these multiple domains are assessed, and the interpretation of performance will take these factors into account (Green, 2000).

RCPM was developed for older people and can be used with people who cannot understand English, or have impairment in verbal language or hearing ability (Raven et al, 2003; Strauss et al, 2006). It is considered more culture fair than the Wechsler test for assessing reasoning ability (O'Leary, Rusch & Guastello, 1991). In addition, the minimal instructions make it an easier test to translate. However, this does not mean that the RCPM is culture free, as even non-verbal tests are culturally biased.

2.4.6 Symbol Digit Modalities Test (SDMT, Smith, 1991):
The SDMT is a widely used measure of neurological impairment (Sheridan, Fitzgerald, Adams, Nigg, Martel, Puttler et al, 2006). The brevity and ease of administration, and its unambiguous scoring, contribute to its widespread use (Strauss et al, 2006). This test took five minutes to complete and did not require verbal language to be completed. The participants were shown a sheet of paper with rows of abstract symbols. Above this row of symbols is a coding key which shows nine abstract symbols with a corresponding number. The participants were instructed to scan the sheet of symbols and assign their respective numbers, as quickly as possible (see Appendix 10 for instructions). The participants are timed and are allocated a time limit

of 90 seconds to match up as many symbols with numbers as possible. The more they completed correctly, the higher the score and the maximum score is 110.

Test-retest reliability was reported by Smith (1991) with an average retest interval of 29 days. Correlations were .90 for the written version and similar values were reported in other studies (Strauss et al, 2006). Uchiyama et al (1994) reported no significant practice effects with the written version following yearly intervals over two years (Uchiyama, D'Elia, Delinger, Selnes, Becker, Wesch et al, 1994). This suggests that the measure would have utility with dementia patients who are tested repeatedly to monitor the progression of the disease.

The SDMT is similar in format to the Wechsler digit symbol/ coding subtest, and correlations ranged from .62 to .91 (Sheridan et al, 2006; Strauss et al, 2006) which suggests good construct validity. However, the SDMT is considered to be more difficult than the Wechsler version. The SDMT also assesses similar constructs to the Trail Making Test (TMT) and PASAT (Royan, Tombaugh & Rees, 2004). However, the PASAT requires spoken language to be completed by the participant and the TMT requires knowledge of the Latin alphabet, therefore reducing their utility with those not fluent in English.

Deficits in attention abilities are a known symptom of dementia (Green, 2000). It is one of the most sensitive tests of attention in neuropsychology (Strauss et al, 2006). However, there are demographic effects on performance on the SDMT. Scores decline with advancing age (Bowler, Sudia, Mergler, Harrison & Cone, 1992) and performance improves with increasing IQ (Uchiyama et al, 1994). However, Sheridan et al reported that age did not significantly affect performance on the SDMT.

2.4.7 **Brixton Test** (Burgess & Shallice, 1997):

Participants were presented with a 56-item stimulus book. Each page showing the same basic array of 10 circles set in two rows of five, with each circle numbered from one to 10. On each page one of the circles was filled with the colour blue. The position of the blue circle changed from one page to the next and the participant was instructed to guess the position of the blue circle on the next page. The blue circle moved according to a simple rule and the participant was expected to reason what the rule was in order to guess where the blue circle will appear on the following page. The changes were governed by a simple rule that changed without warning. The total number of errors across 55 trials was the outcome measure, and a higher score indicated greater impairment of executive functioning. This test is reported to take five minutes to complete (Burgess & Shallice, 1997; Strauss et al, 2006), however in this sample; the approximate time taken to administer this test was 15-20 minutes.

Split-half reliability in healthy individuals is .62 and test retest reliability was marginally better at .70 (Burgess & Shallice, 1997; Strauss et al, 2006). The Brixton test does correlate with other measures of executive functioning, such as the Tower of London (.58) and part B of the Trail Making Test (Tchanturia, Andreluh, Morris, Rabe-Hesketh, Collier, Sanchex et al, 2004).

Duncan et al (1995) showed that performance on executive tests were well predicted by fluid intelligence, leading to suggestions that fluid intelligence may be a more general measure of executive function (Duncan, Burgess & Emslie, 1995). Burgess and Shallice (1997) recommend that tests be used with caution with individuals with relatively low educational attainment, or who pre-morbidly fell below the bottom 15% of the population on measures of general intelligence. In addition, older age is also shown to have a negative

impact on performance (Andres & Van der Linden, 2000; Bielak, Mansueti, Strauss, & Dixon, 2006). However, when fluid intelligence was controlled for, executive performance on no longer correlated with age (Rabbitt & Lowe, 2000). Nevertheless, executive processes are one notable area of cognitive functioning that decline with age (Brennan, Welsh & Fisher, 1997). As a consequence, normative data for a typical ageing population has recently been published for this test (Bielak et al, 2006).

This test did not require verbal responses and the participant could give their answer by pointing to where they expected the circle to be on the next page. The Hayling test is also an appropriate measure of executive function, but is reliant on patients spoken English for administration. There are no studies reporting the relationship between the MMSE and the Brixton test. The cross-cultural equivalence of this measure has not been explored. However, the limited reliance on spoken language from participants may make it easier to translate (Strauss et al, 2006).

2.4.8 Clock Drawing Test (CDT, Tuokko, Hadjistavropoulos, Miller, Horton & Beattie, 1995):
A higher score indicated greater impairment in cognitive functioning. The drawing was scored on the basis of what errors were made in the domains of omissions, perseverations, rotations, misplacements, distortions, substitutions, and additions. The test manual was used to score clock drawings. A maximum of 31 errors could be obtained,, and a higher error score indicated greater impairment.

The CDT is frequently recommended as a screening test for dementia (Libon, Swenson, Barnoski & Sands, 1993; Storey, Rowland, Basic & Conforti, 2001; Richardson & Glass, 2002) There are more than a dozen different versions of the clock-drawing test (Strauss et al, 2006). In addition, scoring systems have been proposed without

agreement on which is best for specific clinical settings (Storey et al, 2001; Strauss et al, 2006). Storey and colleagues used ROC analysis to report the accuracy of five clock scoring methods for dementia detection. The results showed that the Shulman CDT (Shulman, Shedletsky & Silver, 1986) had high inter-rater reliability (.93), intra-rater reliability (.96), and explained the largest Area Under the Curve (.79). However, the Tuokko et al (1995) has similar inter-rater reliability values but provides a comprehensive scoring system and standardised administration materials, which is vital in research settings to minimise confounding variables. This is especially of importance when many different types of errors are present for different reasons, and scoring of drawings become ambiguous. Also, different scoring methods are highly correlated with each other (Tuokko, Hadjistavropoulos, Rae & O'Rouke, 2000; Strauss et al, 2006).

Test-retest reliability coefficient was .70, which indicates that the scores obtained are quite stable (Tuokko, Hadjistavropoulos, Miller & Beattie, 1992). In terms of validity the CDT is not as sensitive in measuring visuospatial, constructional and executive difficulties. However Kurzman (1992) examined the convergent validity of the CDT with neuropsychological measures. Validity correlations ranged from .04 to .27, and validity was higher with dementia patients compared to those with no cognitive difficulties. However, Hadjistavropoulos et al (1991) reported statistically significant correlations with language subtests from the WAIS-R, such as with similarities ($r = -.41$; $P<.001$), visuaspatial subsets such as the Digit Span ($r = -.43$; $P< .001$), and the Block Design ($r = -.57$; $P<.001$). In addition, there were significant correlations with memory subtests (Hadjistavropoulos, Tuokko & Beattie, 1991). Furthermore, errors on the CDT have been found to relate to conceptualization and semantic knowledge (Rouleau, Salmon & Butters, 1996). This suggests that

although the CDT is not a measure of memory, executive functioning or language abilities, the underlying constructs of the CDT tap into these domains, especially in dementia populations where there is global cognitive decline. The CDT also correlates with global measures of cognitive functioning, such as the MMSE with correlations ranging from .41 to .80 (Strauss et al, 2006).

The Tuokko CDT has sensitivity of 92% and specificity of 86% in detecting dementia (Tuokko et al, 1992). Further studies have suggested that although patients with AD and Huntington's Disease have similar quantitative patterns on the CDT, there is a qualitative difference in the type of errors they make (Rouleau, Salmon, Butters, Kennedy & McGuire, 1992). However, the validity of the CDT in terms of its ability to differentiate between different disorders has not been found consistently by research (Suhr, Grace, Allen, Nadler & McKenna, 1998; Strauss et al, 2006).

In comparison to other dementia screening assessments, the CDT is not dependent on verbal responses from patients. In addition, clock faces are familiar to most cultures and the CDT does not require Arabic numerals and participants can use their own culture specific numerical systems to indicate the hours in the clock face (Tuokko et al, 1995). It has been validated in mild AD non-English speaking older people in Taiwan (Chiu, Li, Lin, Chui & Liu, 2008). In addition, Borson and Colleagues (1999) concluded that the CDT was as effective as the MMSE as a first level dementia screen for multilingual samples, even when interpreters were bilingual interpreters were not available (Borson, Brush, Scanlon, Vitaliano, Chen, Cashman et al, 1999).

2.4.9 Colour Trails Test (CTT, D'Elia, Satz, Uchiyama & White, 1996):

The standard Trails Making Test is one of the most frequently administered measures in neuropsychological practice (Lezak et al, 2004). The CTT consisted of two parts. Part 1 and 2 consist of several numbered circles coloured in vivid pink or yellow. All odd numbered circles are coloured in pink, and all even numbered circles are coloured in yellow. In part 2 each number is printed twice, once in pink and once in yellow. In part 1 the participants were instructed to link in ascending order a series of 25 numbers (1-2-3....). Participants were instructed to perform the task as quickly as possible without making errors. In Colour Trails Test 2, the participants were instructed to draw a line between numbered circles, maintaining the sequence of numbers but altering between pink and yellow colours as they proceed (see Appendix 10 for full instructions). The score is time in seconds and a higher score indicated increased impairment.

Tests of attention and information processing are of particular use when assessing dementia (Pachana, Boone, Cummings & Berman, 1996). However, increasing education enhances performance on CTT 2, but less on CTT1 (Mitrushina, Boone, Razani & D'Elia, 2005). The test-retest reliability is moderate (.64) for part 1, and higher (.79) for part 2 (Strauss et al, 2006). The CTT has not been validated against other cognitive tests as much as the TMT (Strauss et al, 2006). However results comparing TMT part A and CTT 1 found that they were correlated significantly ($r = .408$; $P < .05$) (D'Elia et al, 1996).

The CTT was designed as a culture fair analogue to the original Trail Making Test (TMT). The CTT is however free of the language demands of the TMT. Although it has been suggested that acculturation may exert minimal effects on TMT performance (Arnold, Montgomery, Castaneda & Longoria, 1994), its reliance the English alphabet stimulus in part B, severely limit its application in cross

cultural settings. The CTT was developed on the premise that the TMT had limited utility across cultures. It has been used in cross-cultural research to assess driving ability in older people with cognitive difficulties (Elkin-Frankson, Lebowitz, Kapust, Hollis & O'Conner, 2007), with Turkish non-impaired adults to explore the equivalence of the CTT and the TMT (Dubartey, Townes & Mahurin, 2000) and within with cognitive impairment due to HIV-1 n different cultures (Maj, D'Elia, Satz, Janssen, Zaudig, Uchiyama et al, 1993). Furthermore, Cheung et al (2000) that the CTT and TMT were generally fair across Chinese and English (Cheung, Chan, Yip, Cheung & Chan, 2000). However, it further research is required to establish the cross-cultural equivalence of the CTT and TMT (Razani, Burciaga, Madore & Wong, 2007; Strauss et al, 2006).

2.5 Procedure

2.5.1 Recruitment

The lead researcher approached service managers and Health Care Professionals (HCPs) based at older people's Mental Health Services within two NHS Trusts in two large cities in England. Potential participants were identified through a network of professionals working in two memory clinics, four Community Mental Health Teams (CMHT) and two Day Hospitals. The lead researcher provided information about the aims, hypotheses and study methodology. Various methods were used to disseminate this information such as presentations in multi-disciplinary meetings/ specialty meetings, handouts, and communicating through service managers. This provided an opportunity for referrers to meet with the lead researcher, and have questions regarding the study answered. In total the lead researcher made face-to-face contact with 73 mental health professionals over a period of four months. The time taken to present research varied from 30 minutes to an hour. This included a diverse mix of professionals, such as health care assistants,

psychiatrists, nurses, social workers, occupational therapists, speech and language therapists and clinical psychologists.

Attempts at understanding the motivational needs of clinicians can lead to a more accurate perception of how to reach willing collaborators (Young & Dombrowski, 1989). Due to the demands of working in a busy NHS context, the burden on health professionals was minimised. The lead researcher collected all the data and requested only minimal information from referrers. A key incentive, for staff visited, was the need for improved guidlines on deciding on drug eligibility for people with mild, moderate and severe dementia who were unable to complete the MMSE. HCPs were keen for any results to be made available to their teams. This is strongly advocated by Patel and colleagues (2003), who stated that this is neglected by researchers and can have adverse consequences on future attempts to recruit participants in the region (Patel, Doku & Tennakoon, 2003). Hence, results will be disseminated to all collaborators through presentations upon completion of this study.

HPCs' were given copies of invitation letters for participants (see Appendix 3), referrer information sheets (see Appendix 4), and referral forms (see Appendix 5). Initial patient contact was made by HPCs' who informed clients of the study and handed out invitation letters. HPCs' established if patients were interested in hearing more about the study and if so, they sought consent to complete a referral form, with the patients contact details on it, and forwarded this to the lead researcher. The referral forms were either posted to the University address or faxed to a secure site within a NHS trust. The lead researcher made contact after a minimum of 24 hours, via telephone and briefly explained the study and arranged an initial appointment. Memory difficulties are common in dementia, hence participants were encouraged to make a note of the appointment on a

calendar if they had one. Participants who expressed an interest in taking part, but who reported not having a memory aid or a strategy to document appointments, were given the option of having a carer, who was present, to make a record of the appointment and remind them if needed. To minimise confusion or distress due to forgetting, a picture of the lead researcher was included on all documents handed to participants. This enabled them and their carers to know, in advance, who to expect on the visit and also to enable them to verify, for their safety, that the person visiting them was the person who was responsible for the research.

To minimise inconvenience and ensure participants would not have to incur any financial cost for taking part, all data was collected during appointments arranged at participant's homes. During this visit the lead researcher explained the studies objectives and procedures and ensured that the participants met the inclusion criteria. The participants who did not meet the inclusion criteria were excluded at this stage. The lead researcher talked through the Participant Information Sheet (see Appendix 6). A formal capacity to consent assessment (see Appendix 7) was conducted prior to signing the consent form (see Appendix 8). A copy of the consent form was handed to the participant and a copy was included within their clinical notes and their research file.

2.5.2 Data collection

The settings of the assessments varied depending on the mobility needs of the participants. The venues ranged from participant's homes or day hospitals. The lead researcher endeavoured to maintain some consistency across the testing procedure for all involved. Therefore, the following rules were upheld: the lead researcher administered all measures, only the participant and the lead researcher were in the room at the time of the testing, with no

interruptions or distractions, and the room contained a table with two chairs. The lead researcher discussed this with participants and ensured that these conditions could be met prior to booking the appointment. This was done to minimise error variance through confounding variables, such as not having a table to write on. The participants were reminded that they could take a break at any time during the testing.

The lead researcher administered the MMSE according to the standardized administration procedures. The six alternative cognitive tests were also administered. Data was collected in the initial visit for the majority of participants, and over two appointments for three participants. The average total duration of a visit was approximately 2 hours. This involved building rapport with participants, delivering information about the study, answering any questions presented by participants and carers, carrying out a capacity assessment, carrying out cognitive testing and debriefing. The administration of all cognitive tests took approximately one to one and a half hours for each participant (see Appendix 9 for order and instructions for testing).

The measures were scored using published scoring guidelines for each test. All responses were documented on record forms for each test. Each participant had a separate 'Research File' this contained seven test record forms for the measures they had completed, consent form, demographic sheet and scoring sheets. The storage of research files was arranged with the Trent Doctorate of Clinical Psychology course and the data was stored in a secure office at the University of Nottingham. The referral forms were the only sheets which contained participant identifiable information and these were stored separately in another cabinet within the University of Nottingham. Each participant was assigned a number to match their

contact details with their research files, so that results could be sent to them. Participants were given feedback on their performance in the form of a letter outlining their scores on each measure (see Appendix 11). Referrers were also provided with feedback if the participant provided consent (see Appendix 12 for a sample feedback report for a referrer).

3 Extended Results

3.1 Results

Plan of analysis

The overall aim of the statistical analysis was to determine if the RCFT, RCPM, SDMT, Brixton Test, CDT or CTT could provide equivalent information to the MMSE. This was explored using correlation analysis. Therefore, this analysis explored if changes in participants' performance (scores) on the RCFT, RCPM, SDMT, Brixton Test, CDT or CTT were met with similar changes in the scores obtained on the MMSE.

The second aim was to explore whether or not the RCFT, RCPM, SDMT, Brixton Test, CDT or CTT would differentiate between those eligible for treatment with AchI from those who were not. This was explored using Receiver Operating Characteristic (ROC) analysis. ROC analysis was carried out on the measures that had a statistically significant correlation with the MMSE. This was carried out to explore if a cut-off score could be identified on any of the alternative cognitive tests, that was equivalent to the cut-off on the MMSE which determined eligibility for AchI (<20). Only the cut-off score between the mild and moderate range was explored as no participants in this sample obtained a score below 14 points on the MMSE. Therefore, in the absence of participants scoring within the severe range (<9), it was not possible to explore the utility of the alternative cognitive

tests to differentiate between the severe and moderate range on the MMSE.

The mild (>21) and moderate (<20) cut-offs on the MMSE were taken from NICE guidance (NICE, 2007), despite the existence of alternative age and education adjusted cut-offs proposed within the literature. This was because it is the cut-offs recommended by NICE which clinicians are expected to use and any alternative adjusted cut-offs advised in the literature are not used consistently across services. The cut-offs identified using ROC analysis would ascertain what score a patient would need to score on any of the alternative cognitive tests, in order to be eligible for AchI.

3.2 Descriptive statistics of performance on cognitive tests

Table 1: Descriptive statistics of performance on cognitive tests

Test‡	n	Score range on test	Range	Mean	Std. Deviation
MMSE	19	0-30	14-29	22.4	4.2
RCFT VC	19	0-36	2-36	19.8	10.8
RCFT IM	19	0-36	0-23.5	3.1	6
RCFT SP*	19	Time in seconds	108-537	326.3	143.4
RCFT DM	19	0-36	0-22.5	2	5.4
RCFT R	19	0-20	12-20	14.4	4.4
RCPM	20	0-36	5-35	17.5	8.2
SDMT	20	0-115	0-31	14	8.6
Brixton*	18	0-55	11-52	37.2	10.1
CTT 1*	18	Time in seconds	55-370	146.1	83.1
CTT 2*	12	Time in seconds	143-430	251	99.6
CDT*	20	0-31	0-24	10.6	8

‡ RCFT VC, Rey Complex Figure Test Visual Construction; RCFT IM, Rey Complex Figure Test Immediate Memory; RCFT SP, Rey Complex Figure Test Speed of Processing; RCFT D, Rey Complex Figure Test Delayed Memory; RCFT R, Rey Complex Figure Test Recognition Memory; RCPM, Ravens Colour Progressive Matrices; SDMT, Symbol Digit Modalities Test; CDT, Clock Drawing Test; CTT1 & 2, Colour Trails Test 1 & 2.

* Higher score indicates increased levels of impairment.

3.3 Missing data

Pairwise deletion of missing values was used in SPSS version 16. This is appropriate because there were particular tests which participants were unable to complete, such as the Colour Trails Test 2. This was important information because there was a pattern in what tests participants were unable to complete. This demonstrated that certain measures may be too difficult and therefore this may reflect the

appropriateness of using this measure to assess for cognitive functioning in the dementia population. The alternative to this was to use listwise deletion, where the case is excluded from all analyses. This option was not considered appropriate as it would have decreased the sample size for the analyses (Field, 2000). Missing values could have been replaced with the mean value for the sample, but this would have biased scores towards the mean. On certain measures the standard deviation of the mean was so large that a mean score would not have been an accurate reflection of the true score that the participant may have gained on this test.

3.4 Exploring the distribution of scores

Initially the data was explored to establish patterns within it. This was used to see whether the data met the criteria necessary for statistical procedures that were to be conducted for each stage of the analysis. Exploration of the data was carried out using a range of methods on SPSS Version 16. This included pie charts for plotting frequencies and percentages, scatter plots to show relationships to inform correlations and histograms for checking the distribution of scores (Field, 2000; Pallant, 2007).

In order to select the appropriate inferential statistical test, the distribution of scores was explored. Statistical tests make assumptions about how the data is distributed (Field, 2000). Parametric tests assume that scores are normally distributed. When this assumption is violated, non-parametric alternatives are to be used to analyse the data. The analysis of interest here was correlation, so analysing the strength of the relationship between the MMSE and the alternative cognitive tests. To decide which statistical test would be used, the distribution of the data for each cognitive test was explored using histograms (Field, 2000; Pallant, 2007; Dancey & Reidy, 2007). In order to explore if the histograms met the

characteristics of a normal distribution, the following criteria was used:

- That the distribution of scores should be symmetrical about the mean
- The tails should meet the x axis at infinity
- The mean, median and mode are all the same value
- It should be bell-shaped

When exploring the distribution using histograms all measures appeared to show that the distribution of the data was not normally distributed. However, the subjective nature of this type of analysis warranted further analysis. Therefore, Skewness and Kurtosis values were also calculated for each variable to verify what was observed from the histograms (Field, 2000). Table 2 presents the skewness and kurtosis values for each measure.

There is an option to transform scores to normalise them. However, there are no consistent guidelines for what level of skewness is problematic and guidelines for when to transform data are mixed (Norris & Aroian, 2004). However, there can be a loss of data which results in inaccurate interpretation of the data because transformed data becomes a step removed from the original measurement (Field, 2000).

The value for skewness and kurtosis should be 0 in a normal distribution (Field, 2000; Dancey & Reidy, 2007). The further away values are from zero (either positive or negative) the more likely it is that the data are not normally distributed and therefore a non-parametric statistical test should be used. However, Field (2000) suggests that the actual values of skewness and kurtosis are not in themselves informative and instead the values need to be converted into z scores. The z –skew is calculated using (S – 0)/ SE Skew (where S is Skew, and SE is the standard error of the Skew), and

values above 1.96 or below -1.96 are significantly different from that expected by a normal distribution ($p<.05$) and may be problematic (Field, 2000). The z skew and z kurtosis for all cognitive measures are shown in table 2.

Table 2: z skew for all cognitive measures

Measure‡	Skewness	Z skew	Kurtosis	Z Kurtosis
MMSE	-.198	-.4	-.733	1
RCFT VC	-.173	-.3	-.992	1
RCFT I	2.991	5.7	9.445	1
RCFT SP	.095	.2	-1.515	1
RCFT D	3.809	7.3	15.331	1
RCFT R	.756	1.4	-.159	1
RCPM	.449	0.9	-.627	1
SDMT	.643	1.3	.018	1
Brixton	-.780	-1.5	2.586	1
CTT 1	1.518	2.8	2.916	1
CTT 2	.376	0.5	-1.405	1
CDT	.495	1	-1.044	1

‡ RCFT VC, Rey Complex Figure Test Visual Construction; RCFT IM, Rey Complex Figure Test Immediate Memory; RCFT SP, Rey Complex Figure Test Speed of Processing; RCFT D, Rey Complex Figure Test Delayed Memory; RCFT R, Rey Complex Figure Test Recognition Memory; RCPM, Ravens Colour Progressive Matrices; SDMT, Symbol Digit Modalities Test; CDT, Clock Drawing Test; CTT1 & 2, Colour Trails Test 1 & 2.

3.5 Justification for use of parametric and non-parametric analysis

On the basis of z-scores for skewness and kurtosis the distribution of scores on the MMSE, RCFT visual construction trial, RCFT speed of processing trial, RCFT recognition memory, RCPM, SDMT, Brixton test, CTT2 and CDT were normally distributed. However, to verify this

Kolmogorov-Smirnov (K-S) and Shapiro-Wilk statistic was used (see Table 3). A significant value ($p<0.05$) indicates a deviation from normality (Pallant, 2007). On the basis of this analysis the scores on these tests appeared to be normally distributed therefore parametric statistical analysis was used. Pearsons product moment correlation analysis was used to explore the relationship between the above measures and the MMSE.

On the basis of the z-scores for skewness and kurtosis the distribution of scores on the RCFT immediate and delayed memory trial, and the CTT1 were not normally distributed. The K-S test test verified this, with the exception of the CTT1. However, the Shapiro-Wilk value was statistically significant (.018) and evidence suggests that the K-S test can be too stringent for small samples (Pallant, 2007). Therefore non-parametric statistical analysis, specifically spearman correlation was used to explore the relationship between these measures and the MMSE.

Table 3: Kolmogorov-Smirnov & Shapiro-Wilk significance values for each test

Test‡	Kolmogorov-Smirnov (K-S)	Shapiro-Wilk
MMSE	.200	.628
RCFT VC	.200	.305
RCFT SP	.125	.112
RCFT I	.000*	.000*
RCFT D	.000*	.000*
RCFT R	.200	.048*
RCPM	.200	.387
SDMT	.200	.049*
Brixton test	.200	.151
CDT	.200	.100
CTT1	.200	.018*
CTT2	.200	.173

‡ RCFT VC, Rey Complex Figure Test Visual Construction; RCFT IM, Rey Complex Figure Test Immediate Memory; RCFT SP, Rey Complex Figure Test Speed of Processing; RCFT D, Rey Complex Figure Test Delayed Memory; RCFT R, Rey Complex Figure Test Recognition Memory; RCPM, Ravens Colour Progressive Matrices; SDMT, Symbol Digit Modalities Test; CDT, Clock Drawing Test; CTT1 & 2, Colour Trails Test 1 & 2.

* = A significant result ($p<0.05$) indicating that the distribution of scores deviate from normality.

3.6 Effect size

The effect size is an objective and standardised measure of the magnitude of observed effect. The fact that the measure is standardised means that the effect size can be compared across different studies that have measured different variables, or have used different scales of measurement. Many measures of effect size have been proposed. The most common is Cohen's d and Pearson's correlation coefficient (r). The correlation coefficient is one of the most common effect size measures and it is incredible versatile (Field, 2000). A correlation coefficient of 0 means there is no effect, and a value of 1 means that there is a perfect effect, or strong relationship. Cohen (1992) has made some widely accepted suggestions about what constitutes a large or small effect in terms of Persons *r*, these are:

r = 0.1-.23 (small effect)
r = 0.24-.36 (medium effect)
r = 0.37 (large effect)

These guidelines can be used to assess the importance of correlations between the MMSE and the alternative cognitive tests, regardless of the significance of the test statistic (Field, 2000). This is because statistical significance is the probability that that the observed relationship could be due to chance (assuming the null hypothesis is correct), it does not tell you about the strength of any relationship found (Field, 2000).

3.7 Bonferroni corrections

Bonferroni corrections were not carried out prior to conducting the correlation analysis, despite recommendations to do so when multiple comparisons are being made (Cabin & Mitchell, 2000). Bonferroni corrections are employed to reduce Type 1 errors, however they also increase the likelihood of type 2 errors (Perneger, 1998). In addition,

Bonferroni corrections reduce the power of a study below the .80 standard. This intern would require larger sample sizes to detect large effects (Cohen, 1988; Nakagawa, 2004). This is not always practical or possible when carrying out time limited, low-cost research within the NHS. In the context of this research, it would also be difficult to establish which correlations were due to chance and which were not. For instance all the cognitive measures were administered to a sample which was experiencing significant, often global decline in their cognitive functioning. Hence, their performance on each measure could not be entirely isolated from the domain that was assessed in another measure. Furthermore, there is no formal consensus on when Bonferroni corrections should be applied.

3.8 Correlations

Parametric and non-parametric correlation analysis was conducted to explore the relationship between the MMSE and the six alternative cognitive tests. The results showed that there was a statistically significant relationship between the following measures:

The MMSE and the RCFT visual construction trial:

There was a positive linear correlation between these two measure, which suggests that a higher score on the MMSE was associated with a higher score on the RCFT visual constructional task ($r = .609$; $p<0.006$; strong relationship).

The MMSE and the RCFT recognition trial

There was a positive linear correlation between the MMSE and the RCFT recognition trial, which suggests that a higher score on the MMSE was associated with a higher score on the RCFT recognition trial ($r=.496$; $p<0.031$; strong relationship).

The MMSE and the Ravens Colour Progressive Matrices

There was a positive linear correlation between the MMSE and the RCPM, which suggests that higher scores on the MMSE was

associated with higher scores on the RCPM (r= .452; p<0.045; large strong relationship).

The MMSE and the Symbol Digit Modalities test

There was a positive linear correlation between the MMSE and the SDMT. This suggests that higher scores on the MMSE was associated with higher scores on the SDMT (r=.670; p<0.001; strong relationship).

The MMSE and the Colour Trails Test 1

There was a negative linear correlation between the MMSE and Colour Trails Test 1. This suggests that higher scores on the MMSE was associated with lower time to complete the CTT1 (Spearman's rho=-.576; p<0.012; strong relationship).

The MMSE and the Clock Drawing Test

There was a negative linear correlation between the MMSE and the CDT. This suggests that a higher score on the MMSE was associated with lower error rates on the CDT (r=-.577; p<0.008; strong relationship).

3.9 Receiver Operating Characteristic (ROC) analysis

When considering the predictive accuracy of a test, statistics derived from the ROC analysis are the preferred indices of predictive accuracy and effect size (Swets, Dawes, & Monahan, 2000; Harris, 2003; linden, 2006; Craig, Browne, Stringer, & Hogue 2008). ROC curves are valuable tools for the assessment of the accuracy of a test by comparing it with a definitive 'gold standard' (reference standard) test (Obuchowski & McClish, 1997). The 'gold standard' test will specify a cut-off point which distinguishes between normal values (negative cases) from abnormal values (positive cases). Thus, indicating absence and presence of disease. If a less extreme cut-off is used, more patients are indicated as positive cases, thus improving the sensitivity of the test (i.e. the probability of rightfully concluding the disease is present in diseased patients), but at the expense of

deteriorating specificity (i.e. the probability of rightfully concluding the disease is absent in healthy patients). ROC curves are used to describe the possible combinations of sensitivity and specificity, depending on the cut-off point that is chosen (Hout, 2003).

In this study the 'gold standard' test was the MMSE. The accuracy of the alternative cognitive tests (index tests) was compared to MMSE in terms of identifying those who are eligible for treatment with AchI and those who are not.

3.9.1 Classifier performance

Generally, both the sensitivity and specificity of a test need to be known in order to assess its usefulness for a diagnosis (Fawcett, 2006). When selecting a cut-off point the trade off between sensitivity and specificity was considered. This was as follows:

- If the threshold for identifying those eligible for AchI, from the index test, is lowered then the number of false positives increases (the percentage of participants who were not eligible for AchI who were incorrectly classified as being eligible for AchI)
- If the threshold for identifying those eligible for AchI, from the cognitive tests, is heightened then the number of false negatives or misses increases (the percentage of participants who are incorrectly classified as not being eligible for AchI).

A perfect measure would have 100% sensitivity and 100% specificity, thereby correctly identifying everyone and never misclassifying people (Bewick, Cheek, & Ball, 2004). In reality few measures are that accurate (Linden, 2006). The cut-off identified from each of the index tests needed to balance high specificity (>80%) with the least acceptable rate of false positives (>60%). However, there is no clear standard set for what percentage of sensitivity and specificity is

acceptable. The standards used in this research have also been used in a previous studies (Barr, 1997; Blake, McKinney, Treece, Lee & Lincoln, 2002; Bewick et al, 2004; Lepeleire, Heyrman, Baro & Buntinx, 2005; Linden, 2006; Mitchell, 2009).

A limitation of ROC analysis is that the predictive values of the cognitive tests are highly sensitive to the prevalence rate of the observed outcome in the population being evaluated (Altman & Bland, 1994; Linden, 2006). When the sample has a high prevalence of the outcome the Positive Predictive Value increases (PPV), however the Negative Predictive Value (NPV) decreases. Conversely, when the prevalence of positive cases (those who are eligible for AchI) in the sample is low, the PPV decreases and NPV increases. The prevalence of participants who were eligible for AchI in this sample was 40%, therefore, the predictive accuracy of all the measures evaluated in this study was lower, due to the unequal prevalence rates. In addition, due to missing data the prevalence rate of those eligible for AchI for each ROC analysis varied according to the index test being evaluated. The prevalence of those eligible for AchI was calculated for each measure using the following metric: (A+B)/ (A+B+C+D) x 100%.

Cut-offs were identified for all the measures that correlated with the MMSE and these will be presented individually below. Once an appropriate cut-off was identified based on a balance between sensitivity and specificity, the variables were transformed with these cut-offs, within SPSS, using cross tabulations. A kappa analysis was conducted to establish the rate of true negative and false positives identified using these cut offs. Given each measure, there are four possible outcomes. If the instance is positive and it is classified as positive it is counted as a true positive; if it is classified as negative, it is counted as a false negative. If an instance is negative and it is

classified as negative, it is counted as a true negative, if it is classified as a positive, it is counted as false positive. Given each measure and the test set, a two-by-two contingency table can be constructed representing the dispositions of the set of instances (Fawcett, 2006) see Table 4. Once this data has been completed, it is possible to calculate the false positive rate, the true positive rate, the sensitivity, specificity, and the overall positive predictive value of the cut-offs on the cognitive tests, using the metrics presented in Table 5.

Table 4: Contingency table

	Reference standard test		
Index test	**Positive**	**Negative**	**Total**
Positive	True positives A	False positives B	A + B
Negative	False Negatives C	True negatives D	C+D
Total	A + C	B + D	A+B+C+D

Table 5: Metrics

	Metrics
False positive rate	B/ Total negatives
True positive rate	A/ Total positives
Sensitivity	A/ (A+C) x 100
Specificity	D/ (B + D) x 100
Positive predictive Value(PPV)	A/ (A + B)
Negative Predictive Value (NPV)	D / (C + D)
Discriminant Ability	Sensitivity + Specificity / 2

3.9.2 Area under the curve (AUC)

A ROC analysis plots the tests true positive rate (sensitivity) against its false negative rate (1-specificity) and is constructed by estimating the sensitivity and specificity of each test for each of the participants test score. This produces a line of data points across a graph making up the "curve". This graph is a technique for visualising, organising and selecting classifiers based on their performance (Fawcett, 2006). The AUC is a popular summary measure of the accuracy of a test. It serves as an index to describe the discriminatory property of a test, so one does not have to rely solely on visual inspection to determine how well the test performs (Bewick et al, 2004; Linden, 2006). An AUC of 0.5 is a random, an AUC between 0.5 and 0.7 represents moderate, between 0.9 and 1 represents high accuracy, and an AUC of 1 would represent the ideal test (Fischer, Bachmann & Jaeschke, 2003). However, the full AUC has been criticised because it is the

function of both the sensitivity and specificity, therefore the AUC represents the entire range of error rates and gives equal weight to all false positive rates. The volume under the ROC surface of 1/6 corresponds to a test without discriminatory power, and the value of 1 indicates a perfect test.

The AUC value of two or more tests can be used to make comparisons of their predictive accuracy. If one test has a higher AUC value than another, this suggests that it has better predictive value and can be selected. However, caution must be taken when doing comparisons between two ROC volumes, because it is not possible to establish if there is a statistically significant difference between the AUC of two or more different measures, without appropriate computer software (Stephen, Wesseling, Schink & Jung, 2003; Chi & Zhou, 2008; Erkel & Pattynama, 2008).

3.9.3 Positive and Negative Predictive Values

The positive predictive value (PPV) of a test is the probability that a patient has a positive outcome given that they have a positive test result. This is in contrast to sensitivity, which is the probability that a patient has a positive test result given that they have a positive outcome. Similarly, the negative predictive value (NPV) is the probability that a patient has a negative outcome given that they have a negative test result, in contrast to specificity, which is the probability that a patient has a negative test result given that they have a negative outcome. The PPV and NPV were calculated for each measure and the metrics are presented in Table 5.

3.10 ROC analysis for Rey Complex Figure Test – Visual Construction subtest

The prevalence of those eligible for AchI in this analysis was 42% (8 positive & 11 negative). The best cut-off score was selected (see table 6). A score equal or below 20.5 (<20.5) would identify

participants who were eligible for treatment with Achi, with good sensitivity (87%) and adequate specificity (64%). A lower score on this test infers greater impairment in visual constructional and perceptual ability. The new cut-offs identified 11 positive cases.

Table 6: Coordinates of the ROC Curve for RCFT Visual construction

Positive if Less Than or Equal To	Sensitivity	1 - Specificity
1.0000	.000	.000
2.2500	.000	.091
5.7500	.250	.091
10.0000	.375	.091
12.7500	.375	.182
14.7500	.500	.182
16.0000	.625	.182
18.0000	.625	.273
19.2500	.625	.364
19.7500	.750	.364
20.2500	**.875**	**.364**
22.7500	.875	.455
26.5000	.875	.545
28.7500	.875	.636
30.7500	1.000	.636
32.5000	1.000	.727
34.5000	1.000	.909
37.0000	1.000	1.000

In order to calculate the PPV and NPV of this cut-off, the scores for RCFT construction were redefined within SPSS v16 into binary classifiers. Therefore, a score of <20.5 was assigned a value of one (eligible for AchI) and a score equal or above 21 was assigned a value of two (not eligible for AchI). These were cross tabulated with the MMSE; 'eligible' for AchI and 'not eligible' for AchI. To measure the agreement and calculate the metrics presented in Table 5, Kappa measure of agreement is used (see Table 7). Cohen's kappa

measures the agreement between the evaluations of two raters when both are rating the same object. A value of 1 indicates perfect agreement. A value of 0 indicates that agreement is no better than chance. A kappa coefficient of .5 for represents moderate agreement, above .7 represents good agreement, and above .8 represents very good agreement (Pallant, 2007). The agreement between the MMSE and RCFT visual construction was moderate (kappa= -.511; p<.026).

Table 7: Contingency table showing positive and negative predictive values of cut-off of <20.5 on RCFT Visual construction

	MMSE		
RCFT-Visual construction	**Eligible for Aricept (<20)**	**Not eligible for Aricept (>21)**	**Total (n)**
Positives <20.5	True positives 7 (sensitivity = 87.5%)	False positives 4 (35.4%)	11
Negatives >21	False Negatives 1 (12.5%)	True negatives 7 (Specificity = 63.6%)	8
Total (n)	8	11	19

The cut-off (<20.5) identified 11 positive cases, (participants who were eligible for AchI) and identified eight negative cases (participants who were not eligible for treatment with AchI). The PPV of the RCFT visual construction trial is 7/ (7+ 4) = 0.636 (64%), and the NPV is 7/ (1 + 7) = 0.875 (87%). Therefore, 64% of participants scored <20.5, and were within the moderate range on the MMSE, and thus were correctly classified as being eligible for AchI. 87% of the participants' who scored above 21, also scored above the cut off for treatment with AchI.

The discriminant ability of the cognitive tests can be calculated to summarise the performance of the test (see table 5 for the metric). This tells us how much information the cognitive test provides compared to the MMSE, which is assumed to provide perfect (100%) information. The discriminant ability of the RCFT visual construction trial is: (87% + 64%) / 2 = 75%. The AUC = .750 (CI= .524, .976), suggesting moderate accuracy (see Figure 1 and Table 8).

Figure 1: ROC curve for Rey Complex Figure Test- Visual construction

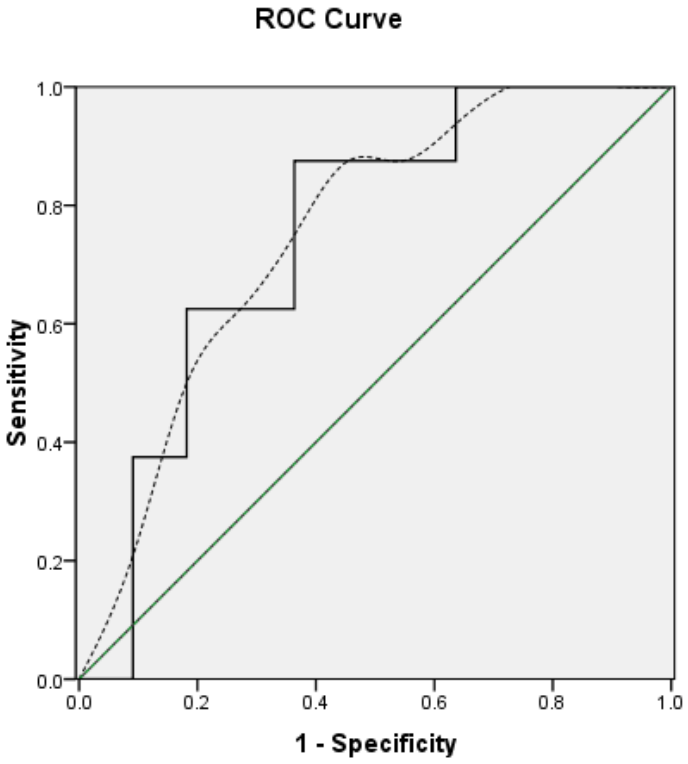


Table 8: Area Under the Curve RCFT Visual construction

AUC	Standard Error	Significance (p<.05)	95% Confidence Interval	
			Lower Bound	Upper Bound
.750	.115	.069	.524	.976

3.11 ROC analysis for Rey Complex Figure Test – Recognition memory subtest

The prevalence of those eligible for AchI in this analysis was 42% (8 positive & 11 negative). The cut-off score selected was <14.5 (a score of 14.5 or below), which identified participants who were eligible for treatment with AchI, with good sensitivity (87%) and good specificity (72%) (see Table 9). This cut off identified 11 positive cases. A lower score on this test infers greater impairment in visual recognition memory.

Table 9: Coordinates of the ROC Curve for RCFT Recognition memory

Positive if Less Than or Equal To	Sensitivity	1 - Specificity
11.0000	.000	.000
12.5000	.375	.091
13.5000	.625	.182
14.5000	**.875**	**.273**
15.5000	.875	.455
16.5000	.875	.636
18.5000	1.000	.818
21.0000	1.000	1.000

The kappa measure of agreement was moderate (kappa= -.592; p<.009). The PPV = 7/ (7 + 3) x 100=70%. The NPV= 8/ (1 + 8) x 100 = 89% (see Table 10). Therefore, 70% of participants scored <14.5, and were within the moderate range on the MMSE, and thus were correctly classified as being eligible for AchI. 89% of the participants' who scored above 21 also scored above the cut off for treatment with AchI.

Table 10: Contingency table showing positive and negative predictive values of cut-off of <14.5 on RCFT Recognition memory

	MMSE		
RCFT-recognition memory	**Eligible for Aricept (<20)**	**Not eligible for Aricept (>21)**	**Total (n)**
Positives <14.5	True positives 7 (sensitivity = 87.5%)	False positives 3 (27.2%)	11
Negatives >15	False Negatives 1 (12.5%)	True negatives 8 (Specificity = 72.7%)	8
Total (n)	8	11	19

The discriminant ability of the RCFT recognition memory trial is: (87% + 72%) / 2 = 79.5%. The AUC = .801 (CI= .590 to 1.012) which suggests moderate to high accuracy (see Figure 2 and Table 11).

Figure 2: ROC curve for RCFT- Recognition memory

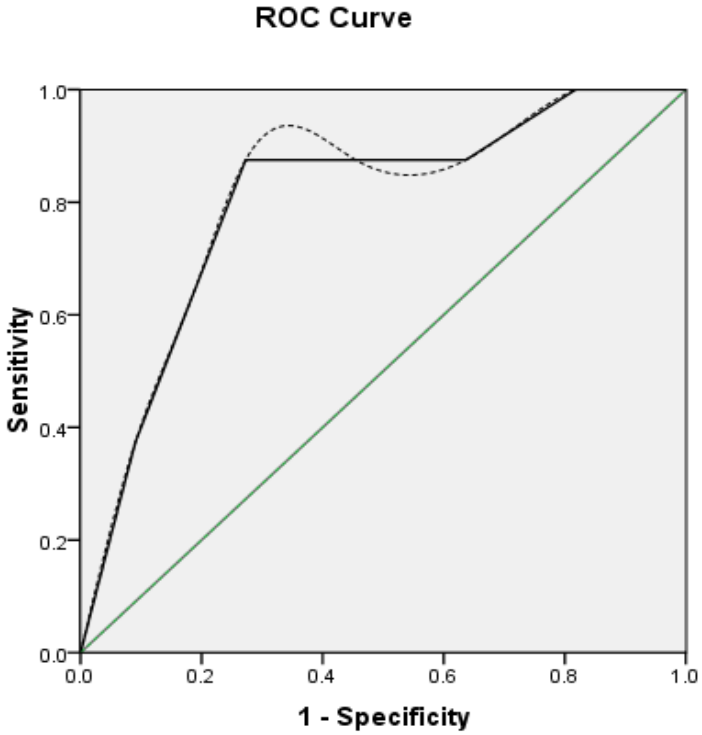


Table 11: Area Under the Curve RCFT Recognition memory

AUC	Standard error	Significance (p<.05)	95% Confidence Interval	
			Lower Bound	Upper Bound
.801	.108	.029	.590	1.012

3.12 ROC analysis for Ravens Colour Progressive Matrices

The AUC value for the RCPM was .573 (see Figure 3 and Table 12). This suggested that the measure was no better than chance at discriminating between positive and negative case, despite a statistically significant correlation (r=.452; p<.05) with the MMSE. Subsequently, it was not possible to identify a cut-off score, which had acceptable sensitivity or specificity (see Table 13).

Figure 3: ROC curve for RCPM

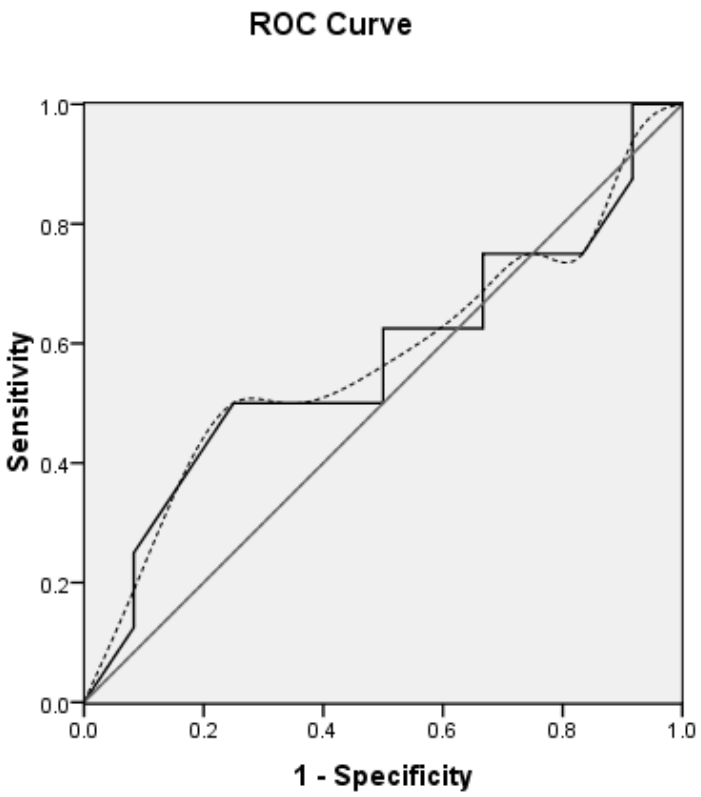


Table 12: Area Under the Curve for the RCPM

Area	Standard error	Significance (p<.05)	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.573	.140	.589	.298	.848

Table 13: Coordinates of the ROC Curve for RCPM

Positive if Less Than or Equal To	Sensitivity	1 - Specificity
5.0000	.000	.000
8.0000	.125	.083
10.5000	.250	.083
12.5000	.500	.250
14.5000	.500	.333
16.0000	.500	.500
18.5000	.625	.500
20.5000	.625	.667
21.5000	.750	.667
24.0000	.750	.750
26.5000	.750	.833
28.5000	.875	.917
32.5000	1.000	.917
36.0000	1.000	1.000

3.13 ROC analysis for Symbol Digit Modalities Test

The prevalence of those eligible for AchI in this analysis was 40% (8 positive & 12 negative). The cut-off score selected was <11 (a score of 11 or below), which identified participants who were eligible for treatment with AchI, with good sensitivity (75%) and adequate specificity (66%) (see Table 14). This cut off identified 10 positive cases. A lower score on this test infers greater impairment in global cognitive functioning.

Table 14: Coordinates of the ROC curve for SDMT

Positive if Less Than or Equal To	Sensitivity	1 – Specificity
-1.0000	.000	.000
.5000	.250	.167
1.5000	.375	.167
3.0000	.375	.250
6.5000	.500	.250
10.0000	.625	.250
11.5000	**.750**	**.333**
12.5000	.875	.417
14.5000	.875	.583
17.0000	.875	.750
19.0000	.875	.833
25.0000	1.000	.833
30.5000	1.000	.917
32.0000	1.000	1.000

The kappa measure of agreement was moderate (kappa= -.400; p<.068). The PPV = 6/ (6 + 4) x 100=60%. The NPV= 8/ (2 + 8) x 100 = 80% (see Table 15). Therefore, 60% of participants scored <11, and were within the moderate range on the MMSE, and thus were correctly classified as being eligible for AchI. 89% of the participants' who scored above 21 also scored above the cut off for treatment with AchI.

Table 15: Contingency table showing positive and negative predictive values of cut-off of <11 on SDMT

	MMSE		
SDMT	**Eligible for Aricept (<20)**	**Not eligible for Aricept (>21)**	**Total (n)**
Positives <11	True positives 6 (sensitivity = 75%)	False positives 4 (33.3%)	10
Negatives >12	False Negatives 2 (25%)	True negatives 8 (Specificity = 66.6%)	10
Total (n)	8	12	20

The discriminant ability of the SDMT trial is: (75% + 66.6%) / 2 = 70.8%. The AUC= .708 (CI= .469 to .947), which suggests moderate accuracy to high accuracy (see Figure 4 and Table 16).

Figure 4: ROC curve for SDMT

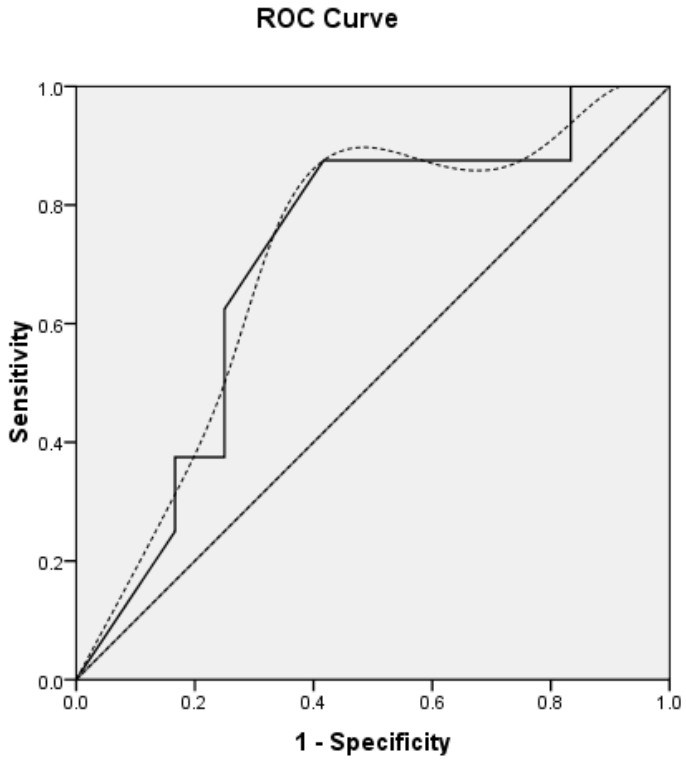


Table 16 Area under the curve

Area	Standard error	Significance (p<.05)	95% Confidence Interval	
			Lower Bound	Upper Bound
.708	.122	.123	.469	.947

3.14 ROC analysis for Colour Trails Test 1

The prevalence of those eligible for AchI in this analysis was 39% (7 positive & 11 negative). On the Colour Trails Test 1 a cut-off of >114.5 seconds (a score of 114.5 or higher) would identify participants who were eligible for treatment with AchI, with good sensitivity (86%) and adequate specificity (64%) (see table 17). A higher score on this test infers greater impairment in attention and executive functioning.

Table 17: Coordinates of the ROC curve for CTT1

Positive if Greater Than or Equal To	Sensitivity	1 - Specificity
54.0000	1.000	1.000
56.0000	1.000	.909
65.0000	1.000	.818
75.5000	1.000	.727
81.0000	1.000	.636
88.5000	.857	.636
95.0000	.857	.545
100.5000	.857	.455
114.5000	**.857**	**.364**
133.0000	.714	.364
144.5000	.714	.273
149.0000	.571	.273
153.0000	.571	.182
165.0000	.571	.091
176.0000	.571	.000
206.5000	.429	.000
238.0000	.286	.000
305.5000	.143	.000
371.0000	.000	.000

The kappa measure of agreement was moderate (kappa= .458; p<.040). The PPV = 6/ (6 + 4) x 100=60%. The NPV = 7/ (1 + 7) x 100 = 88% (see Table 18). Therefore, 60% of participants scored >114.5 seconds, and were within the moderate range on the MMSE, and thus were correctly classified as being eligible for AchI. 88% of the participants' who scored below 115 seconds, also scored above the cut off for treatment with AchI.

Table 18: Contingency table showing positive and negative predictive values of cut-off of >114.5 on the CTT1

	MMSE		
CTT1 (seconds)	**Eligible for Aricept (<20)**	**Not eligible for Aricept (>21)**	**Total (n)**
Positives >114.5	True positives 6 (sensitivity = 86%)	False positives 4 (36%)	10
Negatives <115	False Negatives 1 (14%)	True negatives 7 (Specificity = 64%)	8
Total (n)	7	11	18

The discriminant ability of the CTT1 is: (86% + 64%) / 2 = 75%. The AUC =.818 (CI= .610-1.027), which suggests moderate to high accuracy (see Figure 5 and Table 19).

Figure 5: ROC curve for Colour Trails Test 1

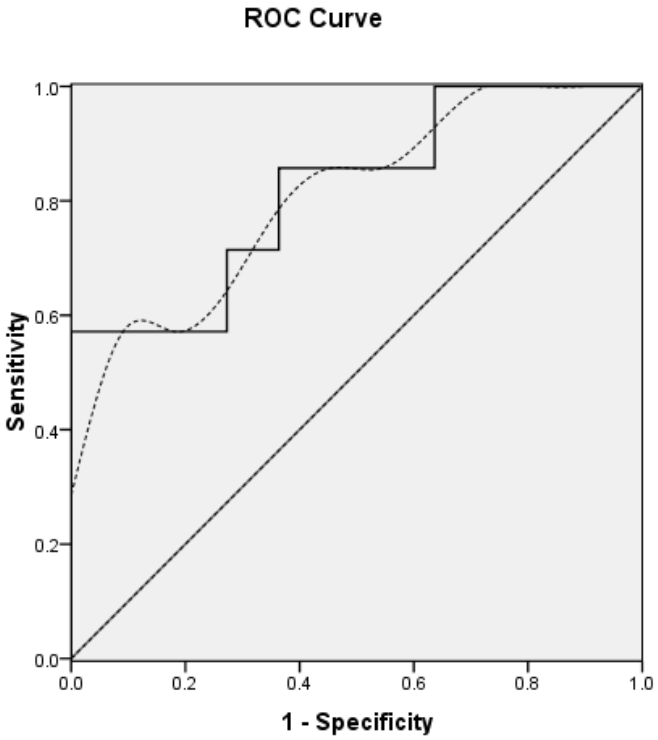


Table 19: Area Under the Curve for the Colour Trails Test 1

Area	Standard error	Significance (p<.05)	95% Confidence Interval	
			Lower Bound	Upper Bound
.818	.106	.026	.610	1.027

3.15 ROC analysis for the Clock Drawing Test

On the Clock Drawing Test a cut-off of >9 errors (a score of 9 or above) would identify participants who were eligible for treatment with AchI, with good sensitivity (75%) and adequate specificity (75%) (see Table 20). A higher score on this test infers greater impairment.

Table 20: Coordinates of the ROC curve for CDT

Positive if Greater Than or Equal To	Sensitivity	1 - Specificity
-1.0000	1.000	1.000
1.0000	1.000	.917
2.5000	.875	.917
3.5000	.875	.750
4.5000	.750	.750
5.5000	.750	.667
6.5000	.750	.500
7.5000	.750	.417
8.5000	.750	.333
9.5000	**.750**	**.250**
12.0000	.625	.250
14.5000	.625	.167
15.5000	.500	.083
18.5000	.375	.083
21.5000	.375	.000
22.5000	.250	.000
23.5000	.125	.000
25.0000	.000	.000

The kappa measure of agreement was moderate (kappa= -.400; p<.068). The PPV = 6/ (6 + 4) x 100=60%. The NPV= 8/ (2 + 8) x 100 = 80% (see Table 21). Therefore, 60% of participants made >9 errors, and were within the moderate range on the MMSE, and thus were correctly classified as being eligible for AchI. 80% of the participants' who made <8 errors, also scored above the cut off for treatment with AchI.

Table 21: Contingency table showing positive and negative predictive values of cut-off of >9 on the CDT

	MMSE		
CDT (errors)	**Eligible for Aricept (<20)**	**Not eligible for Aricept (>21)**	**Total (n)**
Positives >9	True positives 6 (sensitivity = 75%)	False positives 4 (33%)	10
Negatives <8	False Negatives 2 (25%)	True negatives 8 (Specificity = 67%)	10
Total (n)	8	12	20

The discriminant ability of the CDT is: (75% + 67%) / 2 = 71%. The AUC= .734 (CI= .479 to .990), this suggests modest accuracy (see Figure 6 and Table 22).

Figure 6: ROC Curve for the Clock Drawing Test

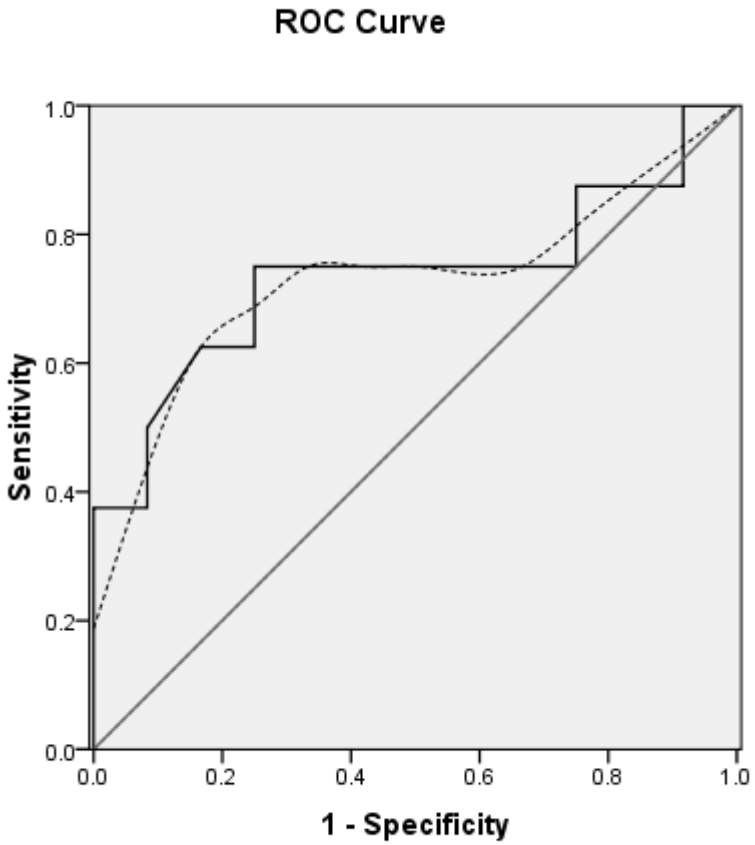


Table 22: Area Under the Curve for the Clock Drawing Test

Area	Standard error	Significance (p<.05)	95% Confidence Interval	
			Lower Bound	Upper Bound
.734	.130	.083	.479	.990

4. Extended Discussion

4.1 Utility of alternative cognitive tests

It is possible to use each of the index tests together to assess for cognitive functioning, or to select individual tests. In order to decide which cognitive test to use, clinicians will need to consider the advantages and disadvantages of using a specific test based on the test properties and the client's experience. Dementia patients experience global decline in cognitive functioning (Green, 2000), therefore it is important to select measures which are tolerated well by this client group. In terms of the Brixton test of executive functioning, Burgess and Shallace (1997) reported that it was tolerated well by older people with dementia. However, observations of participant's behaviour during this research suggested that caution should be taken when administering this test with people with dementia. This is because the test involved presenting the participants with 55 items and there was a lack of a discontinue rule. This meant that all 55 test items needed to be administered according to standardised procedures. This was despite the fact that participants had difficulty comprehending the task and had already responded incorrectly to over 30 test items. Furthermore, when the blue circle did not appear in the position they had expected, they became aware that they were making mistakes. Subsequently, participants who had insight into their difficulties reported their frustration at being confronted with their difficulties in this way. Therefore, further research exploring the possibility of including a discontinue rule would improve the utility of this test with people with dementia.

An additional challenge to selecting reliable and valid cognitive tests for people with dementia is the difficulty in isolating a cognitive domain within an assessment. For instance, due to the prevalence of memory impairment, it was difficult to assess whether performance

on the Ravan's Colour Progressive Matrices or the Brixton test was due entirely to impairment in the respective cognitive domain, or if it was due to forgetting of instructions. Some participants may continue to complete a task implicitly, without following any explicit rules or instructions. This is difficult to control for, which reinforces the need to carry out global assessment of cognitive functioning in clinical practice (Lezak et al, 2004).

Finally, there is a tension between selecting tests where performance is judged on the basis of time (i.e. RCFT speed of processing, SDMT and CTT) and tests that do not have a discontinue rule (i.e. RCPM and Brixton Test). It is suggested that timed test are biased against older people, who have naturally occuring slowed psychomotor and motor processing (Lezak et al, 2004). There are also cultural differences in terms of how much emphasis people place on 'working quickly' or 'getting it right'. Even when participants were told to work as quickly as possible, this did not appear to impact on their performance. However, some older people with dementia experience difficulty with perseveration. Hence tests which do not have specific standardised procedures for discontinuation will limit their utility in this client group.

4.2 Drug treatment practices

Cerejeira and Mukaetova-Ladinska (2007) highlighted that a problem with NICE guidance is that the application of the MMSE might become a substitute for clinical decisions rather than acting as a useful instrument in the clinical-decision making process. They suggested that by rigidly following the MMSE score clinicians would miss the complexity involved in making a therapeutic decision. It was also noted that there was a lack of empirical data about how many clinicians are using AchIs for mild dementia, and how beneficial they are in routine clinical settings. This study found that 80% of those

who were not eligible for treatment with an AchI (i.e. those who scored 21 or over on the MMSE) were still being prescribed AchI. However, given that the participants in this sample were already being prescribed AchI, it is possible that at the time of prescription the participants did have a score within the moderate range on the MMSE. Therefore, their MMSE scores could have improved due to the effect of the medication. However, evidence from RCTs report a maximum of 1 point change in MMSE scores following treatment (Birks & Harvey, 2006). Furthermore, the test-retest reliability of the MMSE is poor (Davey & Jamieson, 2004). Hence it is possible that variations in assessment practices could have lead to participants having different scores when being prescribed the medication, compared to when being selected for this study.

Nonetheless, it would not be surprising if those scoring above the cut-off were being prescribed AchI, as this has been observed in clinical practice (Doyle, 2001). However, there are no adverse implications of this for patients because AchI has been approved and is safe at a maximum dose of 10mg for people with AD or mixed AD and VaD of mild and severe level (Birks & Harvey, 2006). They are merely not deemed to be cost effective or clinically effective for each of these groups (NICE, 2007).

One explanation proposed for why AchIs are being prescribed outside of guidelines is that an abundance of guidance may mean that clinicians place less importance in them (Iliffe & Manthorpe, 2007). Furthermore, clinicians may be more flexible because they are taking into account other factors resulting in MMSE score variance, such as age and education (Tombaugh & McIntyre, 1992). Therefore, technically, they may not be prescribing outside the guidelines, rather they are taking other sources of evidence into account when interpreting scores. However, this can pose some inequalities in prescription practices depending on how flexible a clinician is in

interpreting the guidance, which can cause further confusion for patients and carers (O'Brien & Ballard, 2001). Furthermore, Connelly and Bullock (2003) suggested that many memory specialists involved in the care of dementia sufferers do not feel that guidelines will largely influence clinical practice.

4.3 Implications for clinical practice

The lack of guidelines for people who do not speak English or who are non-verbal within the NICE guidelines presents unacceptable inequalities in assessment practices for these groups. Aside from being unethical, they are illegal. Although there is widespread controversy surrounding the guidelines, the thrust of the argument against the guidelines has so far been about the unfair restriction placed on people with mild dementia and ageism due to the cost effectiveness analysis (Syrett, 2007). Indeed, when NICE were taken to court by the pharmaceutical company, they did not highlight the discrimination against those who were unable to speak English, people with learning disabilities or people with sensory difficulties. This does not mean to say that clinicians were not aware that the MMSE was not administrable with these groups, it merely highlights that minority groups are not only marginalised within NICE guidance, but also by the bodies and organisations that are disputing the validity of NICE's decision. Hence, although the number of RCTs with people of severe and mild dementia has increased to disprove the NICE guidance, a similar empirical drive towards improving assessment processes for marginalised groups so far has not materialised. Subsequently, this study has significant implications for clinical practice, by providing some evidence in an area where progress is of critical importance.

The index tests that were selected for this research can aid in clinical practice from the earliest stages of the assessment process. The

measures were selected specifically so they were easy to administer and were not time consuming. In busy GP practices, where clients are seen for brief consultations, a brief measure of cognitive functioning can assist them to monitor changes in those who are non-verbal (Glasser, 1993). The CDT and SDMT would be ideal in this setting as they require minimum training in the administration of tests and are easy to score. GPs can monitor whether their patients are scoring within the 'moderate' range on these tests and then make appropriate referrals to memory clinics or older people's Mental Health Services for further assessment. Psychiatrists and prescribing clinicians in turn can use this information, alongside their own assessments to decide if it would be appropriate to prescribe AchI. This would be an improvement in assessment procedures for people who are non-verbal, as it would include an objective assessment measure.

Furthermore, clinicians who work within memory clinics within secondary Mental Health services often refer clients to clinical psychologists or neuropsychologists for neuropsychological assessment. This is to aid in the process of diagnosis and also to monitor change overtime. When routine neuropsychological tests are not appropriate, the alternative tests can be used for those who are non-verbal.

However, cognitive tests selected for this study are more sensitive to cognitive dysfunction compared to the MMSE. Therefore, whilst there is a ceiling effect on the MMSE, the alternative cognitive tests presented with a floor effect. The cut-off scores (which are equivalent <21 on the MMSE) for all the index tests are within the impaired range. Given that the mean MMSE score in this sample was 22, participants were more representative of those at the higher level end of the moderate range. This can be problematic because it means that those who score 10-15 on the moderate range on the MMSE may

floor the alternative cognitive tests. This may certainly be the case for the RCFT recognition trial as the cut-off score selected was <14.5, but a score of 12 can be obtained when a client does not recall a single item presented. Likewise the cut-off score selected for the SDMT was <14, however it is possible to obtain a score of up to 115 on this test. In the absence of data from people within the severe range or moderately severe range, it is difficult to know if the index tests selected for this study would also be appropriate for these groups.

Furthermore, the results of this research should be viewed as preliminary findings only, and further research is needed to validate these findings before clinicians can begin using them to decide on treatment eligibility. Promising as the results may be, they need to be understood in the context of the limitations of the study. For instance, there are factors related to the sample which warrant caution when interpreting the findings. A clinical sample which was already being treated with an AchI was used in this study. This is a significant limitation because it means that these two variables could have confounded the results. This is because it is unclear whether the correlations between the MMSE and the alternative cognitive tests were due to a true relationship between the measures in terms of the cognitive domains that were being measured, or due to the effect of the diagnosis and/ or the drug treatment. Subsequently, this study does not provide evidence that the correlations or the cut-off scores would be the same if a non-clinical sample was used, or for those who are not being treated with an AchI.

Given that dementia is a complex disease and clinical population is heterogeneous (Rockwood et al, 2007; Robillard, 2007), the correlations could be unstable. This is because variations in the impact of the disease across the dementia population could mean that the same correlations (Table 1) may not be found in a repeated

study. Although it could also be argued that it is unlikely that the medication would have had a drastic impact on the MMSE scores (Birks & Harvey, 2006), this cannot be said for the alternative cognitive tests. Previous research in this area has shown that neuropsychological tests, such as the Symbol Digit Modalities Test, have demonstrated better outcomes, even when the MMSE score has not changed (Salloway et al, 2004).

This does not invalidate the findings of this research. It does provide justification for further research efforts in this area, specifically using the index tests selected in this study. This would expand on the findings of this study. The limitations could be rectified by administering the seven cognitive tests with a non-clinical sample. In addition, a study using a clinical sample which is not being treated with an AchI would also provide further information about what effect the medication has on the relationship between the selected cognitive tests and the MMSE.

Finally, there is no evidence that these measures are equivalent across cultural groups. Therefore the outcome of this study will not lead to any direct benefits for non-English speakers. Although the measures may result in benefits for people who are non-verbal in the future, it is important to differentiate between those with expressive and receptive language difficulties. Individuals who experience difficulty understanding spoken language would still experience difficulty completing these assessments, because all instructions are delivered using spoken language.

4.4 Additional limitations of study

Diagnosis

There are a number of methodological issues which can arise when conducting research with people with dementia, when recruiting from

different NHS sites, and when limited by resources. The background literature and a period of pre-study consultation with health professionals highlighted potential challenges and these were taken into consideration when planning this study. There is considerable variability in cognitive and functional ability within each dementia type illness (Crystal et al, 1988; Rockwood et al, 2007). Although 'gold standard' diagnostic criteria have been recommended by various bodies for each dementia type illness, there is considerable variability within and across services of diagnostic procedures that are being used (Cerejeira & Mukaetova-Ladinska, 2007). There is evidence to suggest that whilst there is good inter-rater reliability and appropriate levels of sensitivity and specificity, the diagnostic procedures reported within research achieve higher levels of inter-rater reliability than can be found in clinical practice (Crystal et al, 1988). In a resource limited NHS it is not always possible for clinicians to fulfil all of the rigorous criteria of standardised diagnostic procedures, such as detailed laboratory, neurological and neuropsychological examinations. Whilst this is imperative within the context of scientific research to make comparisons across studies, clinicians have to make best use of the resources available to them.

Subsequently, this study did not set specific diagnostic criteria for referrers, apart from identifying participants with a diagnosis of a dementia type illness. Given that this research was time-limited, strict criteria may have compromised the rate of recruitment. However, this imposed restriction on the methodology as data was not collected to validate that the participants had been classified correctly, on the basis of the gold standard classifications systems specified for each dementia type illness (Rockwood et al, 2007). This will make it more difficult to make comparisons with other studies, or research which seeks to expand on this study.

Previous studies which have carried out research using an NHS sample have incorporated extensive diagnostic procedures within their study methodology. This included standardised diagnostic clinical interviews, carer interviews, neuropsychological assessments, detailed review of clinical notes, neurological/ laboratory examinations and assessments of activities of daily living (Robillard, 2007). However, these studies had multiple research personnel, whilst the data collection for the current study was limited to one researcher. If this study was to be repeated, it would be improved by a more rigorous methodology, which should include the collection of detailed information about the sample as described above.

This study recruited participants with any dementia type illness, despite the fact that the NICE guidelines are specific to patients with AD. Consultations with three psychiatrists, highlighted that recruiters may experience difficulty selecting pure cases of a specific dementia type illness because of the high rates of mixed presentations, which is also documented in the literature (Chen, 2004). Furthermore, clinicians also reported that patients with mixed presentations, overlapping symptoms or a moderate score on the MMSE were being treated with AchI, especially if AD was judged to be the dominant disease (NICE, 2007). This is despite the fact that NICE guidance specifically state that AchI should only be prescribed to patients with AD (NICE, 2007). However, the guidelines do state that patients with a mixed presentation can be prescribed Aricept if AD is judged to be the predominant disease (NICE, 2007). In addition, difficulties regarding differential diagnosis can result in uncertainty in providing a definitive pure diagnosis, which is perhaps contributing to the variability in prescribing practices (Doyle, 2001). Including only people with a pure diagnosis of AD would have lead to a lowered recruitment rate and also excluded people for whom the guidance is also relevant (NICE, 2007). In addition, since this research was

interested in how scores on the MMSE would correlate with alternative cognitive tests, and that even pure cases of AD are not a homogenous sample in terms of their cognitive functioning, this would not confound the results of the study. Furthermore, this study was not aiming to provide a profile of how people with AD perform on cognitive tests.

4.5 Alternative cognitive tests

When selecting tests for this study, I was limited by what tests were available and most commonly employed by clinicians. Furthermore, there are multiple cognitive tests which do not require spoken language to be completed; however it was not practical to include any additional measures or to carry out a comprehensive critical review of all cognitive measures prior to conducting the study. Therefore, in the planning stages the 'compendium of neuropsychological tests' (Strauss et al, 2006) was consulted for identification of potential tests, which were then reviewed in more detail. However, given the significant findings of this study, a replication of this study using different measures would be of equal interest.

4.6 Further research

Extension of study

The lead researcher spent considerable time consulting with professionals working within older peoples Mental Health Services for recruitment. This generated interest and formed a collaborative alliance between the researcher and the NHS staff. The small sample size does not reflect a lack of interest from potential referrers. The recruitment rate was lowered due to delays related to changes in clinical research supervision provisions in the area, delays in delivery of test materials from publishers, and time taken for ethical approval from three bodies. Subsequently mental health teams are continuing

to refer participants and therefore data will be collected for a further three months to increase the sample size for an equal prevalence rate of negative and positive cases.

Further statistical analysis

This study did not explore whether a combination of the index tests could increase the predictive accuracy in identifying positive cases for treatment. However, given the unequal prevalence rates in this sample, this may be a redundant analysis to carry out at this stage. However, a replication of this study with a larger sample and equal prevalence rates would warrant further statistical analysis using multiple-logistical regression techniques to explore this more reliably. Based on literature citing that the speed and low cost attributions of the MMSE result in its popularity (Strauss et al, 2006), this could be one model which could be used to inform a regression analysis. The combined cumulative administration time of the SDMT, CDT and CTT1 is 15 minutes. Therefore, the combined predictive accuracy of these measures would be of interest in future analysis. However, one would have to bear in mind that the MMSE is not a gold standard diagnostic test and has psychometric instability. This means that it is unlikely that any test will achieve 100% predictive accuracy, and this may be appropriate as neuropsychological tests are generally more reliable and valid (Lezak et al, 2004).

Administering tests using non-verbal cues

In the planning stages of this study we considered administering the RCPM and CTT using non-verbal cues. It has been found that both these measures can be administered in this way (D'Elia et al, 1996 Raven, 2000; Strauss et al, 2006). However, the lack of standardised guidelines about what gestures to use instead of verbal instructions makes it difficult to administer these tests in this way. Furthermore, any results obtained using this method, in the context of research,

would have made it difficult to establish whether participants' performance was an accurate measure of their cognitive functioning or confounded by their difficulties interpreting non-verbal instructions. The missing data from the CTT2 indicates that non-verbal gestures may cause more confusion within a dementia population. Nonetheless, research exploring the potential of administering the RCPM and CTT using gestures in a dementia population, using quantitative and qualitative methods would expand on this the findings of this study. Qualitative research could explore the participant's experience of completing tests using non-verbal gestures. This is relevant as it may highlight additional factors which can influence performance, such as discomfort, frustration and/ or confusion.

4.7 Critical Reflection

This reflection will begin by outlining the positivist epistemological positioning of this research and why it is necessary for it to be so as it is embedded within the medical model. The strengths and weaknesses of the positivist scientific position will be outlined in relation to how these are manifested in the context of this research. Finally, the post-positivism epistemology is discussed as an appropriate alternative in the context of social research, neuropsychological assessment and the discovery of human experience.

Philosophical position of this research

All research aims to formulate or discover something new. When conducting research, within a clinical setting, using clinical populations to aim to improve assessment practices for medical treatments, you become bound by certain standards or 'philosophical positions' from which your research will need to be based, in order for it to be valued. Indeed, Emanuel and colleague (2000) stated that in

order for research to be valuable it needs to be scientifically valid (Emanuel, Wendler & Grady, 2000). The techniques employed within this research, namely seeking ethical approval or scrutiny over study methodology, testing hypotheses, reviewing psychometric data of tests to learn about their 'reliability' and 'validity', and controlling for confounding variables during assessment are seen as embodying the scientific method or positivist epistemology.

Much contemporary philosophical thought especially that which affects the social sciences, revolves around the question of whether 'science' occupies a special privileged place in human thought about reality (Hughs & Sharrock, 1990). There is ample evidence that many believe it does. For instance, NICE guidelines were developed on the back of evidence gathered from RCTs. RCTs are considered the 'gold standard' in medical research (Altman, 1996). The efficacy of the AchIs reported by carers, which was based largely on carer interviews, did not provide scientifically valid evidence of large enough effect to prove the drugs were effective. Thus, carer reports had minimal impact on guidelines. The use of the MMSE on the other hand as a tool to stratify and measure outcomes gained importance, despite reports of a 1 point change (Birks & Harvey, 2006), which can be due to error variance alone. Thus, this is contrary to what is viewed as being minimal standards in science. The bias towards science is clear, or at least the importance placed on quantifiable data over qualitative data is. Indeed, standards for what makes research with human subjects ethical described by Emanuel and colleagues (2000) stated that "the overarching objective of clinical research is to develop generalisable knowledge to improve health and/ or increase understanding of human biology" (pp. 2701). However, there is some logic to this, especially in the context of a National Health Service and due to the need to provide effective but economical healthcare to the masses. Therefore, research in the context of healthcare provision

does need to produce evidence that is generalisable. However, the notion that evidence from RCTs are generalisable is questionable as they employ restrictive inclusion and exclusion criteria, which make the sample unrepresentative of the population the evidence is to be generalised to.

The initial impetus to carry out this research was about addressing the unacceptable inequalities posed by the scientific method to minority groups. This includes ethnic minority groups and those with disabilities. The 'scientific method' does not cater well for these groups. Hence, they are excluded from RCTs (Britton, 1999) or excluded from essential clinical guidelines (NICE, 2007). This is largely because it is more difficult and costly to control for all the variables that could impact on how they will perform on respected measures. In order to impact on practice in this context, the only evidence which is likely to have any impact is those that are quantifiable and generalisable.

In addition, the drive to generate 'generalisable' results will be motivated in part by the economic gain of this. For instance, there is no financial profit to be made from developing cognitive tests for minority groups. Firstly the scientific enquiry would be time consuming, costly and will only benefit the smallest number of people. The lack of regard for including minority groups within RCTs by pharmaceutical companies for AchI trials is reflective of this. The current market for AchI in the UK is £65 million (Kmietowicz, 2006) and this is without the inclusion of minority groups. Therefore it is possible that making drugs more accessible for these groups is not seen as being worth their while, or necessary.

Neuropsychology can also be described as a positivist disciple due to the emphasis on the scientific method of measuring behavioural correlates of neurological functioning. However, neuropsychology

does acknowledge the multiple sources of score variance, including the environment and culture. Therefore, the combination of assessment modalities such as observations, clinical interviews, carer interviews and a detailed history of developmental and environmental experiences are all equally emphasised (Lezak et al, 2004). The interventions also emphasise rehabilitation and compensation, rather than cure.

There are also important insights gained about human experience using social constructivist methods such as Interpretative Phenomenological Analysis. This technique aims to understand how a person makes sense of their experiences and is less concerned about generalising findings. In the context of this research, valuable insights were gained through observing and engaging with participants. They were able to describe their experience of testing and what informed their performance. For instance, a participant reflected that they were a perfectionist and preferred to get things exactly right or not try at all. Hence, their difficulties with copying the Rey Complex Figure Test were likely to be more about their personality rather than impaired constructional ability. However, when reporting the findings of this study, greater emphasis was placed on findings for which there were documented quantitative evidence. This is because despite individual differences, the quantitative data also provided valuable insights. However, in clinical practice both these sources of evidence can be taken together to make more accurate interpretations of scores. When test scores do not correspond to what is observed in practice, further assessment is required. These ideas are firmly placed within the post positivist epistemology proposed by Karl Popper (Hughs & Sharrock, 1990).

References:

Agarwal, M.R., Ferran, J., Ost, K. & Wilson, K. (1996). Ethics of 'informed consent' in dementia research- the debate continues. *International Journal of Geriatric* Psychiatry*, 11,* 801-806.

Agranovich, A.V. & Puente, A.E. (2007). Do Russian and American normal adults perform similarly on neuropsychological tests? Preliminary findings on the relationship between culture and test performance. *Archives of Clinical Neuropsychology, 22*, 273-282.

Albanese, E., Banerjee, S., Dhanasiri, S., Fernandez, J., Ferri, C., Knapp, M. et al. (2007). *Dementia UK: report to the Alzhiemer's society, Kings College London and London school of economics and political science*. Alzheimer's Society.

Alerton, D.L., & Larson, G.E. (1990). Dimensionality of Raven's Advanced Progressive Matrices items. *Educational and Psychological Measurement*, *50*, 887-900.

Allen, L., Brechin, D., Skilbeck, C., & Fox, R. (2007). The figure copy recall test of the adult memory and information processing battery: inter-rater reliability. *British Journal of Clinical Psychology, 46* (2), 241-245.

Altman, D.G. & Bland, M. (1994). Diagnostic tests 2: predictive values. *British Medical Journal. 309*, 102.

Altman, D.G. (1996). Better reporting of randomised controlled trials: the consort statement. *British Medical Journal*, *313*, 570-571.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental* Disorders (IV-TR) Washington, DC: American Psychiatric Association.

Andres, P., & Van der Linden, M. (2000). Age related differences in supervisory attentional system functions. *Journal of Gerontology: Psychological Science, 55B*, 373-380.

Appelbaum, P. & Grisso, T. (2001). *MacCAT-CR: MacArthur competence assessment tool for clinical research*. Sarasota, Fla: Professional Resource Press.

Ardila, A. (1993). *Culture in our brains: cross-cultural differences in the brain-behaviour relationships*. In Toomela, A. (ed), *cultural guidance in the development of the human mind.* Westport, Connecticut: Ablex publishing, 63-86.

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*, 143-150.

Ardila, A., Lopera, F., Rosselli, M., Moreno, S., Madrigal, L., Arango-Lasprilla, J.C. et al. (2000). Neuropsychological profile of a large kindred with familial Alzheimers disease caused by the E280A single presenilin-1 mutation. *Archives of Clinical Neuropsychology*, *15,* 515-528.

Arnold, B. R., Montgomery, G. T., Castaneda, I., & Longoria, R. (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment*, *1,* 239–248.

Auer, S., Hampel, H., Moller, H-J., & Reisberg, B. (2000). Translations of measurements and scales: opportunities and diversities. *International Geriatrics, 12,* 391-394.

Baird, A.B., Ford, M. & Podell, K. (2007). Ethnic differences in functional and neuropsychological test performance in older adults. *Archives of Clinical Neuropsychology, 22,* 309-318.

Barr, W.B. (1997). Receiver operating characteristic curve analysis of Wechsler memory scale- revised scores in epilepsy surgery candidates. *Psychological Assessment, 9(3)*, 171-176.

Bates, T.C. & Rock, A. (2004). Personality and information processing speed: independent influences on intelligence performance. *Intelligence, 32*, 33-46.

Bewick, V., Cheek, L. & Ball, J. (2004). Statistics review 13: Receiver operating characteristic curves. *Critical Care, 8* (6),508-512.

Bielak, A.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*, 141-149.

Bird, H.R., Canino, G., Stipec, M.R., & Shrout, P. (1987). Use of the Mini-Mental State Examination in a probability sample of a Hispanic population. *Journal Nerv Ment Dis, 175*, 731-737.

Birks, J, & Harvey, RJ. (2006) Donepezil for dementia due to Alzheimer's disease. *Cochrane Database Systematic Review*. 25, (1) Retrieved from: http://mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD001190/pdf_fs.html

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

Boone, K.B., Lesser, I.M., Hill-Gutierrez, E., Berman, N.G., & D'Elia, L.F. (1993). Rey-Osterrieth complex figure performance in healthy, older adults: relationship to age, education, sex, and IQ. *The Clinical Neuropsychologist, 7*, 22-28.

Boone, K.B., Victor, T.L., Wen, J., Razani, J & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22,* 355-365.

Borson, S., Brush, M., Gil, E., Scanlan, J., Vitalino, P., Chen, J. et al. (1999). The Clock Drawing Test: utility for dementia detection in multiethnic elders. *J Gerontol Med Sci, 54(A),* 534– 540.

Bowler, R., Sudia, S., Mergler, D., Harrison, R., & Cone, J, (1992). Comparison of digit symbol and symbol digit modalities tests for assessing neurotoxic exposure, *The Clinical Neuropsychologist, 6,* 103-104.

Brayne, C. & Calloway, P. (1990). The association of education and socioeconomic status with the mini mental state examination and the clinical diagnosis of dementia in elderly people. *Age and Ageing, 19*, 91-96.

Brennan, M., Welsh, M. C. & Fisher, C.B. (1997). Ageing and executive function skills: an examination of a community-dwelling older adult population. *Perceptual and Motor Skills*, *84,* 1187-1197.

Britton, A., McKee, M., Black, N., McPherson, K., Sanderson, C. & Brain, C. (1999). Threats to applicability of randomised controlled trials: exclusion and selective participation. *Journal of Health Services Research Policy*, *4,* 112-121.

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton tests*. Thurston, Suffolk: Thames valley Test Company.

Butt, J., & O'Neil, A. (2004). 'Lets Move On': Black and Minority Ethnic Older People's Views on Research Findings. Joseph Rowntree Foundation, 175, (1). Retrieved from: www.jrf.org.uk/knowledge/findings/socialcare/564.asp

Cabin, R.J., & Mitchell, R.J., (2000). To Bonferroni or not to Bonferroni: when and how are the questions. *ESA Bull, 81*, 246–248.

Carpenter, W.T., Gold, J.M., Lahti, A.C., Queern, C.A., Conley, R.R., Bartko, J.J. et al. (2000). Decisional capacity for informed consent in schizophrenia research. *Arch Gen Psychiatry, 57,* 533-538.

Cerejeira, J. & Mukaetova-Ladinska, E.B. (2007). Treating dementia: will the NICE guidance 2006 change our practice? *Age and Ageing, 36,* 605-606.

Chen, C.P.L.-H. (2004). Transcultural expression of subcortical vascular disease. *Journal of the Neurological Sciences*, *226,* 45-47.

Chervinsky, A., Mitrushina, M., & Satz, P. (1992). Comparison of four methods of scoring the Rey-Osterrieth Complex Figure drawing test on four age groups of normal elderly. *Brain Dysfunction, 5,* 267-287.

Cheung, L.T., Chan, C, Yip, J., Cheung, P. & Chan, C. (2000). Trail making across languages. (Abstract). *Archives of Clinical Neuropsychology, 15,* 702.

Chi, Y. & Zhou, X. (2008). Receiver operating characteristic surfaces in the presence of verification bias. *Applied Statistics*. *57* (1), 1-23.

Chiu, Y., Li, C., Lin, K., Chiu, Y. & Liu, H. (2008). Sensitivity and specificity of the clock drawing test, incorporating Rouleau scoring system, as a screening instrument for questionable and mild dementia: scale development. *International Journal of Nursing Studies, 45,* 75-84.

Church, M. & Watts, S. (2007). Assessment of mental capacity. *Psychiatric Bulletin, 31,* 304-307.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd Ed). New York: Academic Press.

Cohen, J. (1992). "A power primer". *Psychological Bulletin,* 112, 155–159.

Connolly, C.M., & Bullock, R. (2003). Treatment of patients with Alzheimer's disease: A national survey following release of the NICE guidance. *Psychiatric Bulletin*, *27*, 11–3.

Craig, L.A., Browne, K.D., Stringer, I. & Hogue, T.E. (2008). Sexual reconviction rates in the united kingdom and actuarial risk estimates. *Child Abuse & Neglect, 32,* 121-138.

Crystal, H., Dickson, D., Fuld, P., Masur, D., Scott, R., Mehler, M et al (1988). Clinico-pathologic studies in dementia: nondemented subjects with pathologically confirmed Alzheimer's Disease. *Neurology*, *38,* 1682-1687.

Dancey, C.P. & Reidy, J. (2007). *Statistics without maths for psychology* (Fourth Ed). Essex, England: Pearson Education Limited.

Davey, R.J., & Jamieson. S. (2004). The validity of using the mini mental state examination in NICE dementia guidelines. *Journal of Neurology Neurosurgery and Psychiatry,* 75, 343–344.

D'Elia, L.F., Satz, P., Uchiyama, C. L., & White, T. (1996). *Colour Trails Test. Professional manual*. Odessa, FL: Psychological Resources.

Dening, T. (1992). Community psychiatry of old age: a UK perspective. *International Journal of Geriatric Psychiatry. 7*, 757-66.

Department of Health (2000). *The NHS plan: A plan for investment, a plan for reform*. London: HMSO.

Department of Health. (2005). The Mental Capacity Act. London: HMSO.

De Silva, H.A. & Gunatilake, S.B. (2002). Mini Mental State Examination in Sinhalese: a sensitive test to screen for dementia in Sri Lanka. *International Journal of Geriatric Psychiatry, 17,* 134-139.

Doyle, Y. (2001). Equity in the new NHS: hard lessons from implementing a local healthcare policy on denepezil. *British Medical Journal*, *323,* 222-224.

Dugbartey, A.T, Townes, B.D, Mahurin, R.K (2000). Equivalence of the Colour Trails Test and Trail Making Test in Non-native English-Speakers. *Archives of Clinical Neuropsychology*, *15* (5), 425-431.

Duncan, J., Burgess, P. & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia, 33*, 261-268.

Dyer, C. (2006). Nice faces legal challenge over restriction on dementia drugs. *BMJ, 333,* 1085.

Elkin-Frankson, S., Lebowitz, B.K., Kapust, L.R., Hollis, A.M. & O'Conner, M.G. (2007). The use of the colour trails test in the assessment of driver competence: preliminary report of a culture fair instrument. *Archives of Clinical Neuropsychology*. *22,* 631-635.

Emanuel, E.J., Wendler, D. & Grady, C. (2000). What makes clinical research ethical? *Journal of American Medical Association*, *283* (20), 2701-2711.

Eolas. (1999). *Health and Social Care for Older Black and Ethnic Minority Residents of Sefton*. Commissioned by Sefton Health Authority and Sefton Council.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*, 1-11.

Erkel, A.R. & Pattynama, P.M.T. (2008). Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. *European Journal of Radiology, 27*, 88-94.

Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J. & Golding, J.M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity. *Journal of Nervous & Mental Disease*, *174*, 607-614.

Fastenau, P.S, Bennett, J.M., & Denburg, N. L. (1996). Application of psychometric standards to scoring system evaluation: is "new" necessarily "improved"? *Journal of Clinical and Experimental Neuropsychology, 18,* 462-472.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*, 861-874.

Ferraro, R. (2002). Minority and cross-cultural aspects of neuropsychological assessment. Netherlands: Swets & Zeitlinger.

Ferri, C.P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., et al. (2005) Global prevalence of dementia: a Delphi study. *Lancet*, *366*, 2112-2117.

Field, S.J., Jackson, H.J., Hassett, A.M. & Pattison, P. (1995). Ability of the mini-mental state examination to discriminate

diagnostic entities in a psychogeriatric population. *International Journal of geriatric Psychiatry, 10,* 47-53.

Field, A. (2000). *Discovering statistics using SPSS for windows.* London, UK: Sage Publishing.

Fillenbaum, G., Heyman, A., Williams, K., Prosnitz, B., & Burchett, B. (1990) Sensitivity and specificity of standardised screens of cognitive impairment and dementia among elderly black and white community residents. *Journal of Clinical Epidemiology*, *43,* 651-60.

Fischer , J.E., Bachmann, .LM., Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Journal of Intensive Care Medicine, 29*, 1043–1051.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12,* 189-198.

Folstein, M.F., Folstein, S.E. & Fanjiang, G. (2000). *Mini-Mental State Examination: clinical guide*. Florida, USA: Psychological Assessment Resources, Inc.

Fratiglioni, L., Jorm, A.F., Grut, M., Viitanen, M., Holmen, K., Ahlbom, A. et al. (1993). Predicting dementia from the Mini-Mental State Examination in an elderly population: the role of education. *Journal of Clinical Epidemiology*, *46,* 281-287.

Ganguli, M., Ratcliff, G., Chandra, V., Sharma, S., Gilby, J., Pandav, R. et al. (1995). A Hindi version of the MMSE: The development of a cognitive screening instrument for a largely illiterate rural elderly population in India. *International Journal of Geriatric Psychiatry*, *10,* 367-377.

Gibbons, L. E., van Belle, G., Yang, M., Gill, C., Brayne, C., Huppert, F. A. et al.(2002). Cross cultural comparison of the mini-mental state examination in United Kingdom and United States participants with Alzheimer's Disease. *International Journal of Geriatric Psychiatry*, *17,* 723-728.

Gladsjo, J., Schuman, C., Evans, J., Peavy, G., Miller, S., & Heaton, R (1999). Norms for Letter and category Fluency: Demographic corrections for age, education and ethnicity. *Assessment*, *6,* 147-178.

Glasser, M. (1993). Alzheimer's disease and dementing disorders: practices and experiences of rural physicians. American *Journal of Alzheimer's Disease and Other Dementias*, *8* (4), 28-35.

Godley.F. (2006). Editors choice: fear not the heat. *British Medical Journal*. 333.

Green, J. (2000). *Neuropsychological Evaluation of the Older Adult: A Clinician's Guidebook*. London, UK: Academic press.

Grigoletto, F., Zappala, G., Anderson, D. W. & Lebowitz, B.D. (1991). Norms for the mini-mental state examination in a healthy population. *Neurology, 53*(2), 315–20.

Hajistavropoulos, T., Tuokko, H. & Beattie, B.L. (1991). The clock test: construct validity for a multidimensional assessment method. *Canadian Psychology, 32*, 219.

Harris, G. (2003). Men in his category have a 50% likelihood, but what half is he in? *Sexual Abuse: A Journal of Research and Treatment, 15,* 389-392.

Heiat, A., Gross, C. & Krumholz, H. (2002). "Representation of the elderly, women and ethnic minorities in heart failure clinical trials. *Archives of Internal Medicine*, *162,* 1682-1688.

Helms, J.E. (1997). The triple quandary of race, culture, and social in standardized cognitive ability testing. In Flanagan, D.P., Genshaft, J.L., & Harrison, P.L (Eds). Contemporary intellectual assessment: theories, tests, and issues. New York: Guilford Press.

Hinton, J. & Withers, E. (1971). The usefulness of the clinical tests of the sensorium. *British Journal of Psychiatry, 119,* 9-18.

Holzer, C., Tischler, G., Leaf, P., & Myers, J. (1994). An epidemiologic assessment of cognitive impairment in a community population. In Greenley, J.R (Ed.). *Research in community and mental health, 4,* 4, (pp. 3-32). New York: JAI Press.

Hout, W.B. (2003). The area under an ROC curve with limited information. *Medical Decision Making*, *23,* 160-166.

Hughs, J. & Sharrock, W. (1990). *The philosophy of social research*. London: Longman.

Huppert, F.A, Cabelli, S.T, & Matthews, F.E. (2005). Brief cognitive assessment in a UK population sample- distributional properties and the relationship between the MMSE and an extended mental state examination. *British Medical Council Geriatrics. 5,*7. Retrieved from: http://bmc.ub.uni-potsdam.de/1471-2318-5-7/1471-2318-5-7.pdf

Hussain-Gambles, M., Atkin, K. & Leese, B. (2004). Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health and Social Care in the Community*, *12* (5), 382-388.

Ivnik, J. R. (2005). Normative psychology: a professional obligation. *Clinical Psychologist*, *19,* 159-161.

Jefferson, A.L., Consentino, S.A., Ball, S.K., Bogdanoff, B., Kaplan, E., & Libon, D.J. (2002). Errors produced on the Mini-Mental State Examination and neuropsychological test performance in Alzheimer's disease, ischemic vascular dementia, and Parkinson's disease. *Journal of Neuropsychiatry and Clinical Neuroscience, 14,* 311-320.

Jeste, D.V. & Saks, E.S (2006). Decisional capacity in mental illness and substance use disorders: empirical database and policy implications. *Behaviour Sciences and Law*, *24,* 607-628.

Jeste, D.V., Palmer, B.W., Appelbaum, P.S., Golshan, S., Glorioso, D., Dunn, L.B et al (2007). A brief instrument for assessing decisional capacity for clinical research. *Archives of General Psychiatry*, *64* (8), 966-974.

Johnson, M.R.D, Owen, D., Blackburn, C., Rehman. H. & Nazroo, J. (2000). *Black & Minority ethnic groups in England: the second health & lifestyles survey*. London: Health Education Authority.

Jorm, A.F., Scott, R., Cullen, J.S. & Mackinnon, A.J. (1991). Performance of the informant Questionanire on cognitive decline in the elderly (IQCODE) as a screening test for dementia. *Psychological Medicine, 21* (3), 785-790.

Kahle-Wrobleski, K., Corrada, M.M., Li, M. & Kawas, C.H. (2007). Sensitivity and specificity of the Mini-Mental State Examination for identifying dementia in the oldest-old: the 90+ study. *Journal of American Geriatric Society, 55,* 284-289.

Katzman, R., Zhang, M., Qu, O.Y., Wang, Z., Liu, W.T., Yu, E et al. (1988) A Chinese version of the Mini-Mental State Examination: Impact of illiteracy in a Shanghai dementia survey. *Journal of Clinical Epidemiology*. *41,* 971-978.

Kavirajan, H. & Schneider, L.S. (2007). Efficacy and adverse effects of cholinesterase inhibitors and memantine in vascular dementia: a meta-analysis of randomised controlled trials. *Lancet Neurology, 6,* 782-792.

Kim, S.Y & Caine, E.D. (2002). Utility and limits of the mini-mental state examination in evaluating consent capacity in Alzheimer's disease. *Psychiatric Services, 53* (10), 1322-1324.

Kmietowicz, Z. (2006). Experts defend NICE against attack by US politician. *British Medical Journal*, *333* (25), 1087.

Krecic-Shepard, M.E., Park, K., Barnas, C., Slimko, J., Kerwin, D.R. & Schwartz, J.B. (2000). Race and sex influences clearance of nifedipine: results of a population study. *Clinical Pharmacology and Therapeutics*, *68,* 130-142.

Kurzman, D. (1992). *The construct validity of the clock test in normal and dementia adults*. Masters Thesis. Victoria. B.C: University of Victoria.

Larson, E. (1992) The Ni-Hon-Sea project. In Curb, J.D. & Graves, A.B. (1992). Multi-National Epidemiological Studies of Dementia. *Gerontologist, 32* (2), 219.

Lepeleire, J., Heyrman, J., Baro, F. & Buntinx, F. (2005). A combination of tests for the diagnosis of dementia had a significant diagnostic value. *Journal of Clinical Epidemiology, 58,* 217-225.

Lezak, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological Assessment* (4th ed.). New York, Oxford: University Press.

Libon, D. J., Swenson, R.A., Barnoski, E.J., & Sands, L.P. (1993). Clock-drawing as an assessment tool for dementia. *Archives of Neuropsychology*, *8*, 405-415.

lin, J.T. & Zelinski, E.M. (1991). The 12-month reliability of the mini-mental state examination. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3* (3), 427-432.

Linden, A. (2006). Measuring diagnostic and predictive accuracy in disease management: anintroduction to receiver operating characteristic (ROC) analysis. *Journal of Evaluation in Clinical Practice, 12* (2), 132-139.

Lindesay, J., Jagger, C., Mlynik-Szmid, A., Sinorwala, A., Peet, S. & Moledine, F. (1997a). The Mini-Mental State Examination (MMSE) in an elderly immigrant Gujarati population in the United Kingdom. *International Journal of Geriatric Psychiatry, 12* (12), 1155-1167.

Lindesay. J., Jagger, C., Hibbett, M.J., Peet, S.M. & Moledina, F. (1997b). Knowledge, uptake and availability of health and social services among Asian Gujarati and white elderly persons. *Ethnicity Health*, *2*, 59–69.

Lliffe, S. & manthorpe, J. (2007). Dementia: still muddling along?. *British Journal of General Practice*, 606-607.

MacKenzie, D.M., Copp, P., Shaw, R.J. & Goodwin, G.M. (1996). Brief cognitive screening for the elderly: a comparison of the mini-mental state examination (MMSE), abbreviated mental test (AMT) and mental status questionnaire (MSQ). *Psychological Medicine*, *26,* 427-430.

Maj, M., D'Elia, L., Satz, P., Janssen, R., Zaudig, M., Uchiyama, C. Et al. (1993) Evaluation of two new neuropsychological tests designed to minimize cultural bias in the assessment of HIV-1 seropositive persons: A WHO study. *Archives of Clinical Neuropsychology*, *8*, 123-135.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D. & Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under

the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, *34,* 939-944.

Melzer D. (1998). New drug treatment for Alzheimer's disease: lessons from healthcare policy. *British Medical Journal, 316,* 762-764.

Meyers, J., & Meyers, K. (1995). *Rey Complex Figure and Recognition Trial: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Mills, C.J., Ablard, K.E., & Brody, L.E. (1993). The Raven's progressive matrices: Its usefulness for identifying gifted/talented students. *Roeper Review, 15,* 183-186.

Mitchell, A.J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*. *43,* 411-431.

Mitrushina, M.N., Boone, K.B., Razani, J., & D'Elia, E.F. (2005). *Handbook of normative data for neuropsychological assessment.* (2nd ed). New York: Oxford University Press.

Molloy, D.W., Alemayehu, E., Roberts, R. (1991). Reliability of a standardized mini-mental state examination compared with the traditional mini-mental state examination. *American Journal of Psychiatry, 148, 102*–15.

Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioural Ecology, 15* (6), 1044-1045.

National Audit Office (2007). *Improving Services and Support for People with Dementia.* London: National Audit Office.

National Institute for Clinical Excellence. (2007). Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease (amended). Retrieved from: http://www.nice.org.uk/nicemedia/pdf/TA111fullversionamendedSept07.pdf

Nell, V. (1999). *Cross-cultural neuropsychological assessment: theory and practice.* New York: Lawrence Erlbaum Associates.

Nilsson, F.M. (2007). Mini-mental state examination (MMSE)-probably one of the most cited papers in health science. *Acta Psychiatrica Scandinavica*, *116* (2), 156-7.

Norris, A. & Aroian, K. (2004). To transform or not to transform skewed data for psychometric analysis: that is the question. *Nursing Research, 53* (1), 67-71.

O'Brien, J.T., & Ballard, C.G. (2001). Drugs for Alzheimer's disease. Cholinesterase inhibitors have passed NICE's hurdle. *British Medical Journal. 323,*123-124.

Obuchowski, N.A. & McClish, D.K. (1997). Sample size determination for diagnostic accuracy studies involving binormal ROC curve indices. *Statistics in Medicine. 16,* 1529-1542.

O'Connor, D.W., Pollitt, P.A., Hyde, J.B., Fellows, J.L., Miller, N.D., Brook, C.P.B. et al. (1989). The reliability and validity of the Mini-Mental State in a British community survey. *Journal of Psychiatric Research, 23,* 87-96.

Office for National Statistics (2003). *Census 2001 Key statistics for local authorities in England and Wales*. London: The stationary Office.

O'Leary, U-M., Rusch, K.M. & Guastello, S.J. (1991). Estimating age-stratified WAIS-R IQs from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology, 47,* 277-284.

Olsson, R. (2004). Reality comprehension clock test (RCCT): as an instrument for non-verbal cognitive function in a polish population. *Journal of Multicultural Nursing and Health, 4*, Retrieved from: http://findarticles.com/p/artciles/mi_qa3919/is_200410/ai_n9459587

Pachana, N. A., Boone, K. B., Miller, B. L., Cummings, J. L., & Berman, N. (1996). Comparison of neuropsychological functioning in Alzheimer's disease and frontotemporal dementia. *Journal of the International Neuropsychological Society*, *2*, 505–510.

Pallant, J. (2007). *SPSS survival manual* (3ed). Berkshire, UK: Open University Press.

Park, J., Park, Y.N. & Ko, H,J. (1991). Modification of the Mini-Mental State Examination for use with the elderly in a non-western society. Part II: Cut-off points and their diagnostic validities. *International Journal of Geriatric Psychiatry, 6,* 875-882.

Parker, C. & Philip, I. (2006). Screening for cognitive impairment among older people in black and minority ethnic groups. *Age and Ageing*, *33* (5), 447-452.

Patel, M.X., Doku, V. & Tennakoon, K. (2003). Challenges in recruitment of research participants. *Advances in Psychiatric Treatment*, *9,* 229-238.

Pelosi, A.J., McNulty, S.V., & Jackson, G.A. (2006). Role of cholinesterase inhibitors in dementia care needs rethinking. *British Medical Journal*, *333,* 491–493.

Perez-Arce, P. (1999). The influence of culture on cognition. *Archives of Clinical Neuropsychology*, *14* (7), 561-592.

Perneger, T.V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal, 316,* 1236–1238.

Peterson, R.C., Stevens, J.C., Ganguli, M., Tangalos, E.G., Cummings, J.L. & DeKoski, S.T. (2001). Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology 56*: 1133-1142.

Pryse-Phillips, W. (1999). Do we have drugs for dementia? *Archives of Neurology, 56,* 735-737.

Purandare, N., Luthra, V., Swarbrick, C. & burns, A. (2007). Knowledge of dementia among South Asian (Indian) older people in Manchester, UK. *International Journal of Geriatric Psychiatry*, 22, 777-781.

Rabbit, P. & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research, 63*, 308-316.

Rait, G., Burns, A., Baldwin, R, Morley, M., Chew-Graham & St Leger, S. (2000). Validating screening instruments for cognitive

impairment in older south Asians in the United Kingdom. *International Journal of Geriatric Psychiatry*, *15,* 54-62.

Ramsey, V., Stevens, S., Bryan, K., Binder, J. & Cockle-Hearne, J. (2009). Using the Barnes language assessment with with older ethnic minority groups. Int J Geriatr Psychiatry, **24**, 426-431.

Raschetti, R., Albanese, E., Vanacore, N. & Maggini, M. (2007). Cholinesterase inhibitors in mild cognitive impairment: a systematic review of randomised trials. *PLOS Medicine, 4* (11), 1818-1827.

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, *41,* 1-48.

Raven, J., Raven, J.C., & Court, J.H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview.* San Antonio, TX: Harcourt Assessment.

Razani, J., Burciaga, J., Madore, M. & Wong, J. (2007). Effects of acculturation on tests of attention and information processing in an ethnically diverse group. *Archives of Clinical neuropsychology, 22,* 333-341.

Richardson, A.E. & Glass, J.N (2002). A comparison of scoring protocols on the clock drawing test in relations to ease of us, diagnostic group, and correlations with mini-mental state examination. *American Geriatrics Society, 50,* 169-173.

Robillard, A. (2007). Clinical diagnosis of dementia. *Alzheimer's & Dementia, 3,* 292-298.

Rockwood, K., Bouchard, R.W., Camicioli, R. & Leger, G. (2007). Toward a revision of criteria for the dementias. *Alzheimer's & Dementia, 3,* 428-440.

Rodda, J. & Walker, Z. (2009). Ten years of cholinesterase inhibitors. *International Journal of geriatric Psychiatry, 24,* 437-442.

Rouleau, I., Salmon, D.P., Butters, N., Kennedy, C. & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer's Disease and Huntington's Disease. *Brain and Cognition, 18,* 70-87.

Rouleau, I., Salmon, D.P., Butters, N. (1996). Longitudinal analysis of clock drawing in Alzheimer's Disease patients. *Brain and Cognition, 31,* 17-34.

Royan, J., Tombaugh, T.N. & Rees, L. (2004). The adjusted Paced Serial Attention Test (Adjusted PASAT): thresholds for speed of information processing as a function of stimulus modality and problems complexity. *Archives of Clinical Neuropsychology, 19* (1), 131-143.

Salloway, M.S., Ferris, S., Kluger, A., Goldman, R., Griesing, T., Kumar, D. et al. (2004) Efficacy of donepezil in mild cognitive impairment: a randomized placebo controlled trial. *Neurology, 63,* 651–657.

Salthouse, T.A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology, 84*, 171-179.

Sellal, F., Woltt, V. & Marescaux, C. (2004). The cognitive pattern of vascular dementia and its assessment. *Seminars in Cerebrovascular Disease and Stroke, 4,* 79-86.

Shah, S. & Reichman, W.E. (2006). Treatment of Alzheimer's disease across the spectrum of severity. *Clinical Interventions in Ageing*, *1* (2), 131-142.

Sheridan, L.K., Fitzgerald, H.E., Adams, K.M., Nigg, J.T., Martel, M.M., Puttler, L.I. et al (2006). Normative symbol digit modalities test performance in a community based sample. *Archives of Clinical Neuropsychology, 21,* 23-28.

Shulman, K., Shedletsky, R., & Silver, I. (1986). The challenge of time: clock-drawing and cognitive functioning in the elderly. *International Journal of Geriatric Psychiatry, 1,* 135-140.

Singh, B. & O'Brien, J.T. (2009). When should drug treatments be started for people with dementia? *Maturitas*, *62,* 230-234.

Smith, A. (1991). *Symbol digit modalities test*. Los Angeles, CA: Western Psychological Services.

Starr, J.M. & Lonie, J. (2008). Estimated pre-morbid IQ effects on cognitive and functional outcomes in Alzheimer's Disease: a longitudinal study in a treated cohort. BMC Psychiatry. Retrieved from: http://www.medscape.com/viewarticle/575011

Stephan, C., Wesseling, S., Schink, T. & Jung, K. (2003). Comparison of eight computer programs for receiver-operating characteristics analysis. *Clinical Chemistry, 49 (3),* 433-439.

Storey, J.E., Rowland, J.T.J., Basic, D. & Conforti, D.A. (2001). A comparison of five clock scoring methods using ROC (receiver operating characteristics) curve analysis. *International Journal of Geriatric Psychiatry*, *16,* 394-399.

Storey, J.E., Rowland, J.T.J., Basic, D., & Conforti, D.A. (2002). Accuracy of the Clock Drawing Test for detecting dementia in a multicultural sample of elderly Australian patients. *International Psychogeriatrics, 14*, 259–271.

Strauss, E., Sherman, E.M.S. & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms and commentary* (3rd ed). New York, USA; Oxford University Press.

Stuss, D.T. & Levine, B. (1996). The dementias: nosological and clinical factors related to diagnosis. *Brains & Cognition*, *31,* 99-113.

Sugarman, J., Roter, D., Cain, C., Wallace, R., Schmechel, D. & Welsh-Bohmer, K.A. (2007). Proxies and consent discussions for dementia research. *The American Geriatrics Society*, *55,* 556-561.

Suhr, J., Grace, J., Allen, J., Nadler, J. & McKenna, M. (1998). Quantitative and qualitative performance of stroke versus normal elderly on six clock drawing systems. *Archives of Clinical Neuropsychology, 13* (6), 495-502.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1* (1), 1-26.

Syrett, K. (2007). NICE and judicial review: enforcing accountability for reasonableness through the courts? Eisai limited v. National institute of clinical excellence. *Medical Law Review.* 1-14.

Tang. M., Stern, Y., Marder, K., Bell, K., Gurland, B., Lantigua, R. et al. (1998). The APOE-*e*4 Allele and the risk of Alzheimer Disease among African Americans, Whites and Hispanics. *Journal of American Medical Association, 279,* 751–755.

Tchanturia, K., Andreluh, M.B., Morris, R.G., Rabe-Hesketh, S., Collier, D.A., Sanchex, P. et al. (2004). Cognitive flexibility in anorexia nervosa and bulimia nervosa. *Journal of the International Neuropsychological Society, 10,* 513-520.

Thompson, P., Inglis. F., Findlay. D., Gilchrist. J., McMurdo, M.E.T. (1997). Memory clinic attenders: a review of 150 consecutive patients. *Aging Mental Health*, *1*,181-183.

Tombaugh, T. & McIntyre, N. (1992) The Mini-Mental State Examination: A comprehensive review. *Journal of American Geriatric Society, 40,* 922-935.

Tombaugh, T. N., McDowell, I., Kristjansson, B., & Hubley, A.M. (1996). Mini-Mental State Examination (MMSE) and the modified MMSE (3MS): a psychometric comparison and normative data. *Psychological Assessment, 8* (1), 48-59.

Tuokko. H., Hadjistavropoulos, T., Miller, A., & Beattie, B.L. (1992). The clock test: a sensitive measure to differentiate normal elderly from those with Alzheimer disease. *Journal of American Geriatric Society, 40*, 579–584.

Tuokko. H., Hadjistavropoulos, T., Miller, A., Horton, A. & Beattie, B.L (1995). *The clock test: administration and scoring manual.* New York, USA: Multi-Health Systems Inc.

Tuokko. H., Hadjistavropoulos, T.,Rae, S. & O-Rourke, N. (2000). A comparison of alternative approaches to the scoring of clock drawing. *Archives of Clinical Neuropsychology, 15* (2), 137-148.

Uchiyama, C.L., D'Elia, L.F., Delinger, A.M., Selnes, O.A., Becker, J.T., Wesch, J.E. et al. (1994). Longitudinal comparison of alternative versions of the symbol digit modalities test: issues of form comparability and moderating demographic variable. *The Clinical Neuropsychologist, 8,* 209-218.

Uzzell, B.P., Ponton, M. & Ardila, A. (2007). International handbook of cross-cultural neuropsychology. New Jersey, USA: Lawrence Erlbaum Associates, Inc.

Warner. J., McCarney, R., Griffi, N.M., Hill, K. & Fisher, P. (2008). Participation in dementia research: rates and correlates of capacity to give informed consent. *Journal of Medical Ethics, 34,* 167–170.

Warner, J. & Nomani, E. (2008). Giving consent in dementia research. *The Lancet, 372,* 183-185.

Whitehouse. P.J., Price, D.L., Struble, R.G., Clark, A.W., Coyle, J.T., & Delong, M.R. (1982). Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science*, *215,* 1237–1239.

Whitehouse, P. (1986). The concept of cortical and subcortical dementia: another look. *Annals of Neurology, 19*, 1-6.

Whitehouse P, Maurer K, Ballenger J. (2000). *Concepts of Alzheimer Disease: Biological, Clinical, and Cultural perspectives*. Baltimore, Md: Johns Hopkins Press.

Wilkins, C.H., Wilkins, K.L., Meisel, M., Depke, M., Williams, J. & Edwards, D.F. (2007). Dementia undiagnosed in poor older adults with functional impairment. *The American Geriatric Society, 55,* 1771-1776.

Wind, A.T., Schellevis, F.G., Staveren, G.V., Scholten, R.J.P.M., Jonker, C. & Van Eijk, J.M. (1997). Limitations of the Mini-Mental State Examination in diagnosing dementia in general practice. *International Journal of Geriatric Psychiatry, 12,* 101-108.

Young, C.L. & Dombrowski, M. (1989). Psychosocial influences on research subject recruitment, reenrolment and retention. *Social Work in Health Care, 14,* 43-57.

Zaidel, E., Zaidel, D. & Sperry, R.W. (1981). Left and right intelligence: case studies of ravens progressive matrices following brain bisection and hemidecortication. *Cortex, 17*, 167-186.

APPENDICES

Appendix 1: Instructions to Authors

International Journal of Geriatric Psychiatry

1. AIMS & SCOPE

The rapidly increasing world population of aged people has led to a growing need to focus attention on the problems of mental disorder in late life. The aim of the *International Journal of Geriatric Psychiatry* is to communicate the results of original research in the causes, treatment and care of all forms of mental disorder which affect the elderly. The Journal is of interest to psychiatrists, psychologists, social scientists, nurses and others engaged in therapeutic professions, together with general neurobiological researchers.

The Journal provides an international perspective on the important issue of geriatric psychiatry, and contributions are published from countries throughout the world. Topics covered include epidemiology of mental disorders in old age, clinical aetiological research, post-mortem pathological and neurochemical studies, treatment trials and evaluation of geriatric psychiatry services.

2. MANUSCRIPT CATEGORIES

The *International Journal of Geriatric Psychiatry* invites the following types of submission:

Research Articles

Research Articles are the Journal's primary mode of scientific communication. Peer-review of Research Articles will be handled by the most appropriate Editor. Research Articles must not exceed 3500 words of body text, and are limited to 6 figures/tables.

MANUSCRIPT PREPARATION

Manuscripts must be written in English.

Text should be supplied in a format compatible with Microsoft Word for Windows (PC). Charts and tables are considered textual and should also be supplied in a format compatible with Word. All figures (illustrations, diagrams, photographs) should be supplied in jpg, tiff or eps format.

All manuscripts must be typed in 12pt font and in double space with margins of at least 2.5 cm.

Manuscripts must comply with the word limits defined in section 2, and include:

Title Page

The first page of the manuscript should contain the following information:

- the title of the paper
- a running head not exceeding 50 characters
- 2–6 article keywords and up to 4 key points
- names of authors
- names of the institutions at which the research was conducted
- name, address, telephone and fax number, and email address of corresponding author
- the name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s)
- the word count of the body text

Structured Abstracts

Authors submitting Research and Review Articles should note that structured abstracts (maximum 250 words) are required. The structured abstract should adopt the format: Objective, Methods, Results, Conclusions. (Authors of Reviews may use Design instead of Method.) Abstracts should contain no citation to other published work.

This should in general, but not necessarily, be divided into sections with the headings: Introduction, Methods, Results, Discussion, Conclusion.

Tables and Figures

Tables and figures should not be inserted in the appropriate place in the text but should be included at the end of the paper, each on a separate page.

Tables and figures should be referred to in text as follows: Figure 1, Figure 2; Table 1, Table 2. The place at which a table or figure is to be inserted in the printed text should be indicated clearly on a manuscript. Each table and/or figure must have a legend that explains its purpose without reference to the text.

References

References should be in 'Harvard' format, i.e, names and dates in brackets in the text (Jones, 2000; Smith and Jones, 2001; Jones *et al* ., 2002), and the full reference listed at the end of the paper, in alphabetical order by first author, as follows:

Porteus SD. 1959. *The Maze Tests and Clinical Psychology*. Pacific Books: Palo Alto.

Rabbitt PMA. 1982. How do old people know what to do next? In *Aging and Cognitive Processes* , Craik FIM, Trehub S (eds). Plenum Press: New York; 79–98.

Chou K-L, Chi I. 2004. Combined effect of vision and hearing impairment on depression in elderly Chinese. *Int J Geriatr Psychiatry* 19 : 825–832. DOI: 10.1002/gps.1174

(Titles of periodicals should be abbreviated according to the style used in Index Medicus.)

5. DECLARATION

Conflict of Interest

Authors are responsible for disclosing all financial and personal relationships between themselves and others that might bias their work. To prevent ambiguity, authors must state explicitly whether potential conflicts do or do not exist. Investigators should disclose potential conflicts to study participants and should state in the manuscript whether they have done so. Authors should describe the role of the study sponsor(s), if any, in study design, in the collection, analysis and interpretation of data, in the writing of the report and in the decision to submit the report for publication. If the supporting source had no such involvement, the authors should so state.

Ethics

When reporting experiments on human subjects, indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 1983. Do not use patients' names, initials or hospital numbers, especially in illustrative material. When reporting experiments on animals, indicate whether the institution's or a national research council's guide for, or any national law on, the care and use of laboratory animals was followed. A statement describing explicitly the

ethical background to the studies being reported should be included in all manuscripts in the Materials and Methods section. Ethics committee or institutional review board approval should be stated.

Patients have a right to privacy that should not be infringed without informed consent. Identifying information should not be published in written descriptions, photographs and pedigrees unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication. Informed consent for this purpose requires that the patient be shown the manuscript to be published. Identifying details should be omitted if they are not essential but patient data should never be altered or falsified in an attempt to attain anonymity. Complete anonymity is difficult to achieve and informed consent should be obtained if there is any doubt. For example, masking the eye region in photographs of patients is inadequate protection of anonymity.

Authorship

All persons designated as authors should qualify for authorship and all those who qualify should be listed. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content. One or more authors should take responsibility for the integrity of the work as a whole, from inception to published article. Authorship credit should be based only on 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; 3) final approval of the version to be published. Conditions 1, 2 and 3 must all be met. Acquisition of funding, the collection of data or general supervision of the research group, by themselves, do not justify authorship. All others who contributed to the work who are not authors should be named in the Acknowledgements section.

Appendix 2: NRES ethical approval

NHS

National Research Ethics Service

Leicestershire, Northamptonshire & Rutland Research Ethics Committee 2

1 Standard Court
Park Row
Nottingham
NG1 6GN

Telephone:
Facsimile:

13 February 2009

Miss Sobia T Khan
Trainee Clinical Psychologist
Trent DclinPsy
I-WHO, University of Nottingham
International House, B Floor
Jubilee Campus, Wollaton Road,
Nottingham, NG8 1BB

Dear Miss Khan

Full title of study: **The assessment of dementia severity using non-verbal cognitive tests**
REC reference number: **08/H0402/158**

Thank you for your letter of 06 February 2009, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA). The favourable opinion for the study applies to all sites involved in the research. There is no requirement for other Local Research Ethics Committees to be informed or SSA to be carried out at each site.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission at NHS sites ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Covering Letter		28 November 2008
Protocol	1	22 October 2008
Investigator CV		26 November 2008
Application	AB/132565/1	28 November 2008
Demographic Information	1	26 November 2008
Participant Information Sheet: Referrer Information Sheet	1	26 November 2008
Advertisement	1	26 November 2008
Investigator CV		
MMSE Test		
Scoring Card for Coloured Progressive Matrices		
Rey Complex Figure Test		
Hayling & Brixton Tests		
Results Summary Sheet	1	26 November 2008
Summary/Synopsis: Flowchart	1	26 November 2008
Response to Request for Further Information		06 February 2009
Participant Consent Form	2	06 February 2009
Participant Information Sheet	2	06 February 2009
Letter of invitation to participant	2	06 February 2009
The assessment of participant's capacity to consent	2	06 February 2009

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

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review –guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

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@nres.npsa.nhs.uk.

08/H0402/158	Please quote this number on all correspondence

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

Yours sincerely

[redacted]

Chair / Committee Coordinator

[redacted]

Enclosures: "After ethical review – guidance for researchers" *SL- AR2 for other studies*

[redacted]

Copy to: *R&D office for NHS care organisation at lead site – Derby Mental Health*

Appendix 3: Participant invitation letter

Lead Researcher:
Sobia Khan
Trainee Clinical Psychologist

Picture

Trent Doctorate of Clinical Psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

The assessment of dementia severity using non-verbal cognitive tests

Invitation letter

Dear Sir / Madam,

We would like to invite you to participate in a research study taking place in your area. We would appreciate it if you would take some time to read the following information and share it with others if you like. If any aspect of this information is unclear or if you would like further information please contact the lead researcher on the contact details provided at the end of this letter.

What is the purpose of the study?
This study is interested in the assessment of older people with mental health difficulties as a result of dementia. The purpose of this study is to find alternative methods of assessing people's difficulties and needs, when the usual assessment methods cannot be used.

Dementia is an illness most common in people over the age of 65, but can also affect people under 65. Dementia can cause difficulties with mental abilities such as memory, attention, concentration, vision and reading. This can make it harder for people to remember information and cause them to become confused. It is important for health care professionals to carry out assessments to find out what difficulties the individual is experiencing and what treatment could be helpful for them.

One method of assessing how individual's memory, attention, vision and language are working is to use assessments that test their mental abilities. However, the most commonly used tests are in English and people need to be able to understand and speak English to do them. So, if a person cannot speak English or has difficulty with spoken language it can be difficult to test their mental abilities. The aims of this research are to examine whether alternative tests which do not require you to be able to read, write or speak English can be just as useful.

Why have I been chosen?
You have been chosen because you are over 65 and you can speak English. You have also been chosen because you may be being assessed or treated by a mental health professional because you are experiencing difficulties with mental processes such as memory, attention and concentration.

Do I have to take part?
No, it is up to you whether or not you take part. If you decide to take part you will be given a consent form which you will need to sign if you decide to take part and you will be given a copy of both the signed consent form and an information sheet. If you decide that you would prefer to not take part, this would not affect the standard of care you receive.

What will happen to me if I take part?
If you are happy to take part, your nurse/ doctor will pass your telephone number to the lead researcher, who will call you to arrange an appointment. The researcher will talk through the study with you and you can ask questions. If you choose to take part, we would arrange another appointment to carry out some tests or if you have time we could do them the same day. This will take approximately one hour and fifteen minutes and you will be allowed to take a break at any time. The tests are similar to puzzles involving drawing, remembering objects or carrying out tasks using instructions.

Will my taking part in this study be kept confidential?
Yes. We will follow ethical and legal practice and all information which is collected about you during the course of the research will be kept strictly confidential. The research team will only have access to your research notes, not any medical notes.

Contacts for further information

If you require any further information regarding the study please contact the following people who are part of the research team. If you would prefer independent advice from a body that is not involved in this study, you can contact Alzheimer's Society or the Patient Advice and Liaison Services.

Lead Researcher Sobia Khan (Trainee Clinical Psychologist) University of Nottingham Mobile:	Academic Research Supervisor Dr Roshan Nair (Consultant Clinical Psychologist) University of Nottingham **Tel:**
Clinical Research Supervisor Dr David Connelly: (Consultant Clinical Psychologist) Highbury Hospital **Tel:**	
Independent agencies who you can contact for further advice about taking part in this study:	
Patient Advice and Liaison Services (PALS) PALS Duncan Macmillan House Porchester Road, Nottingham NG3 6AA Freephone Helpline: 0800 0153367	Alzheimer's Society 7 Mansfield Rd Nottingham, NG1 3FB 0115 934 8468 **Tel:** 0115 934 8468

Thank you for taking the time to read this invitation letter and considering taking part in this study.

Appendix 4: Referrer information sheet

Trent Doctorate of Clinical Psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

The assessment of dementia severity using non-verbal cognitive tests

Referrer Information sheet

Dear Colleague,

You are being invited to take part in a research study by assisting with the recruitment of participants. It is hoped that this information will provide you with the necessary details so to make an informed decision about approaching your patient caseloads. Please feel free to contact the lead researcher for further information, or if you are interested in any part of this research and should like to discuss it further.

Who is conducting the research?

I am a second year trainee clinical psychologist currently training on the Trent Doctorate of Clinical Psychology. I am conducting this research as part of my professional training and qualification.

Background

The National Institute of Clinical Excellence has recommended for the Mini-Mental State Examination (MMSE) to be used to assess if a patient with dementia as a result of Alzheimer's Disease may be eligible for treatment with Cholinesterase inhibitors such as Aricept. The guidelines state, those who score within the moderate range (10-20 points) on the MMSE can be prescribed Cholinesterase inhibitors, however, the MMSE has been designed for patients who are able to speak English. The NICE acknowledge that those who are unable to complete the MMSE due to language barriers, should not be discriminated against, but do not provide any additional guidance for clinicians. It is important to research into alternative forms of cognitive assessment for older people with dementia to ensure that individuals from diverse cultural and ethnic backgrounds and people who are non-verbal have equal access to services within the NHS.

Aims of research

The aim of this study is to determine if alternative cognitive assessments, that do not rely on participants English speaking ability, can provide equivalent information to the MMSE and differentiate those eligible for treatment with cholinesterase inhibitors such as Aricept, from those who are not.

The hypotheses are:

1. There will be a significant correlation between participant's total scores on the MMSE and their total scores on the Rey-Osterrith Complex Figure Test (ROCF), Colour Trails Test (CTT) Raven's Colour Progressive Matrices (RCPM) the Clock Drawing Test (CDT), the Brixton test and the Symbol Digit Modalities Test (SDMT).

2. The participant's performance on the ROCF, CCT, RCPM, CDT, the Brixton test and the SDMT will differentiate those eligible for treatment of Aricept (10-20 moderate score on MMSE) from those who are not.

What will my patients have to do?

This will be a within groups design. A total of 37 participants are needed in this study to complete all seven of the above tests either at their homes or at service settings. The testing will take approximately 1 hour and 15 minutes and participants will be allowed a break at any time during testing.

Inclusion Criteria	Exclusion Criteria
Older adults (65+)	Non-English speakers
Capacity to consent	Bilinguals (unless English is their first language)
Cognitive Difficulties – Patients who are experiencing significant cognitive difficulties and are seeking assessment, but it is too soon to confirm a diagnosis.	Diagnosis of a mood disorder which is likely to be impacting on their cognitive abilities.
Diagnosis of a Dementia type illness – any type including mix presentations.	Difficulties with fine motor skill (unable to write/draw/hold a pencil)
All English speaking	Blind
	Deaf

What do I have to do?

I would appreciate it if you could identify potential participants from your case loads, which meet the inclusion and exclusion criteria, and find out if they may be interested in taking part in this research project. I will provide you with invitation letters to hand out to your patients and to find out if they are happy for you to forward their contact details to myself. I will make contact to arrange an informal discussion or visit to talk through the participant information sheet and gain informed consent.

Who has reviewed the research?

This study has been reviewed and given favourable opinion by a Research Ethics Committee in your local area. In addition to this, research supervisors based at the University of Nottingham will monitor the research and provide supervision to the lead researcher. The clinical aspects of the research will be supervised by a qualified clinical psychologist working in your area.

What will happen to the results of the research?

The participants will receive a short report outlining their performance on tests if they wish. If the participant provides consent, this information will also be made available to referrers.

The results of the research will be written up as part of a doctorate thesis. It is hoped the research will lead to publication and that any results will be disseminated in an appropriate format to the staff teams who support the recruitment of participants, the older people's mental health services in Nottinghamshire, and the ethics and R & D departments.

A flow chart has been attached to the end of this information sheet to illustrate the process of this research project. If you would like further information please do not hesitate to contact myself or Dr Roshan Das Nair (Consultant Clinical Psychologist/ Academic research supervisor) on the numbers provided at the end of this information. Thank you for taking the time to read this information sheet and considering taking part in this research.

Yours sincerely

Sobia Khan
Lead Researcher/ Trainee Clinical Psychologist

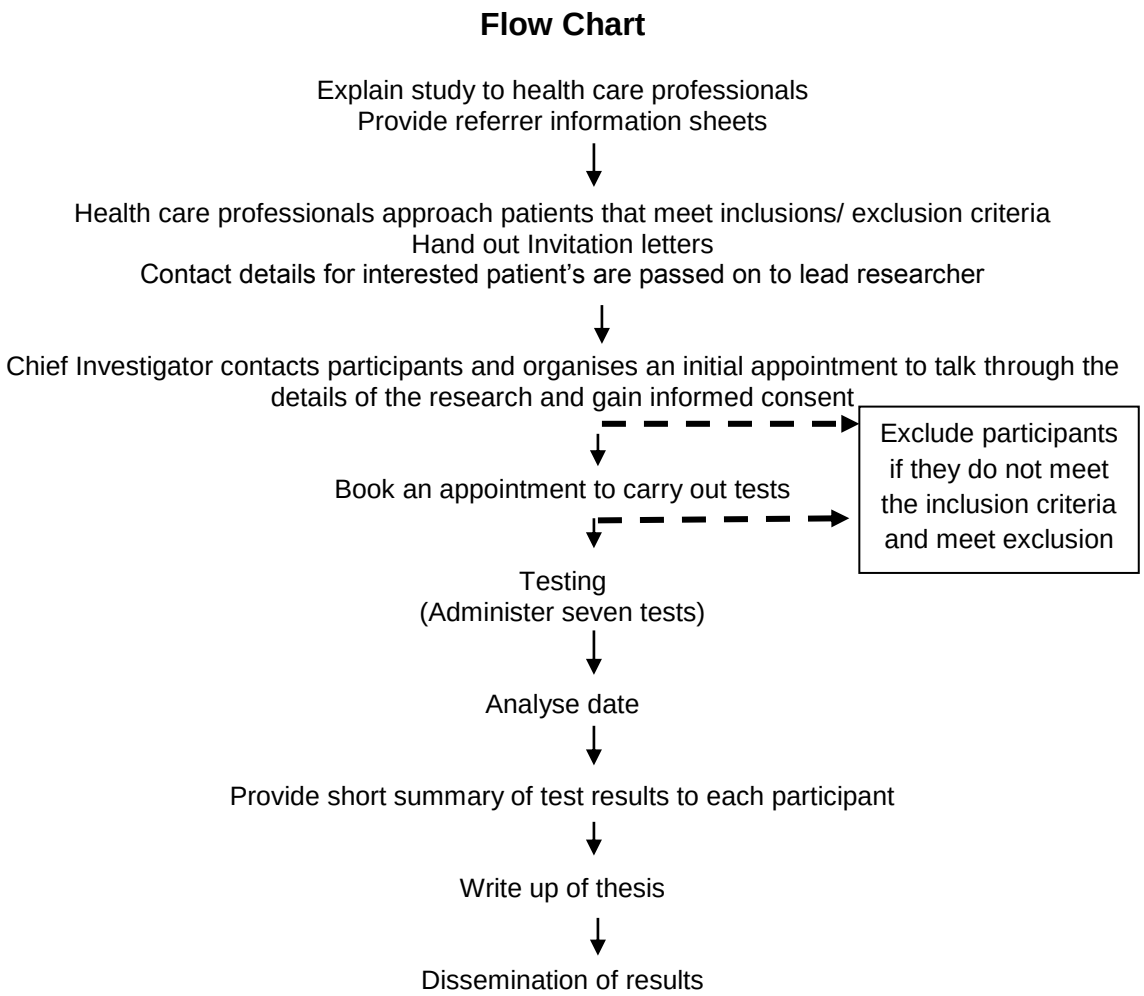
Contact details:

Lead Researcher
Sobia Khan (Trainee Clinical Psychologist)
University of Nottingham
Tel:
Mobile:

Clinical Research Supervisor
Contact details

Academic Research Supervisor
Contact details

The assessment of dementia severity using non-verbal cognitive tests



Appendix 5: Referral form

The assessment of dementia severity using non-verbal cognitive tests

<table>
<tr><td>Lead Researcher: Sobia Khan
Trainee Clinical Psychologist
Tel: 07974211006 or 07805535541</td><td rowspan="2">Contact Details:
Trent Doctorate of Clinical psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

Tel: 0115 8466646</td></tr>
<tr><td>Clinical research supervisor:
Dr David Connelly
Consultant Clinical Psychologist
Tel: 0115 9770000
Fax: 01158542202</td></tr>
<tr><td colspan="2">Thank you for taking the time to talk to your patients about this research project. Please hand the Invitation letters to your patients when you first tell them about the research. If they have expressed an interest in this study or would like further information, please note their contact details below. I would also appreciate it if you could make a note of your patient's most recent MMSE score. Please ensure your patients consent to their information being passed on to me.</td></tr>
<tr><td colspan="2">Participant details</td></tr>
<tr><td>Participant Name:</td><td>Carer/ significant other:</td></tr>
<tr><td>Address:</td><td>Address (if different):</td></tr>
<tr><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td>Post code: Tel:</td><td>Post code: Tel:</td></tr>
<tr><td>Last MMSE score:</td><td>Date MMSE was administered:</td></tr>
<tr><td>Clinician/ Referrer name:

Job title:
Date of referral:</td><td>Place of work:

Contact Number:</td></tr>
<tr><td colspan="2">Thank you for the above information and your contribution to this research. Please post this form to the lead researcher at the address provided above or fax it to David Connelly at the number above.</td></tr>
</table>

Appendix 6: Participant information sheet

Lead Researcher:
Sobia Khan
Trainee Clinical Psychologist

Picture

Trent Doctorate of Clinical Psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

The assessment of dementia severity using non-verbal cognitive tests

Participant Information Sheet

Dear Sir / Madam,
Thank you for considering taking part in this research. Please read the information in this letter for details of the research. If you would like further information please contact the lead researcher on the contact details at the end of this information.

What is the purpose of the study?

This study is interested in the assessment of older people with mental health difficulties as a result of dementia. The purpose of this study is to find alternative methods of assessing people's difficulties and needs, when the usual assessment methods cannot be used.

Dementia is an illness most common in people over the age of 65, but can also affect people under 65. Dementia can cause difficulties with mental abilities such as memory, attention, concentration, vision and reading. This can make it harder for people to remember information and cause them to become confused. It is important for health care professionals to carry out assessments to find out what difficulties the individual is experiencing and what treatment could be helpful for them.

One method of assessing how individual's memory, attention, vision and language are working is to use assessments that test the individual's mental abilities. However, the most commonly used tests

are in English and people need to be able to understand and speak English to do them. So, if a person cannot speak English or has difficulty with spoken language it can be difficult to test their mental abilities. The aims of this research are to examine whether alternative tests which do not require you to be able to read, write or speak English can be just as useful.

Why have I been chosen?
You have been chosen because you are over 65 and you can speak English. You have also been chosen because you may be being assessed or treated by a mental health professional because you are experiencing difficulties with mental processes such as memory, attention and concentration.

Do I have to take part?
No, it is up to you whether or not you take part. If you decide to take part you will be given this information sheet to keep and be asked to sign a consent form to show you have agreed to take. You are free to withdraw at any time and without giving reason. This would not affect the standard of care you receive.

What will happen to me if I take part?
You will complete some mental ability tests with the lead researcher. This will take approximately one hour and fifteen minutes. The tests are similar to puzzles, involving drawing, remembering objects or carrying out tasks using instructions. We can do these tests at your own home or you could come to a clinic near you. We would need to make sure that there were no distractions so you can concentrate and you have a table that you can lean on to write.

What are the possible disadvantages or risks of taking part?
We do not anticipate that there will be many disadvantages or risks in taking part in this project. The tests cannot harm you, however, some people can feel stressed when they are being tested. There is no need to worry about your performance on the tests. You are not expected to know all the answers because some of the tests are designed to get progressively more difficult. We are not asking you to do the tests to find out if you have an illness and your results on the tests will not affect the care you are receiving.

If you have any concerns during testing you can talk to the lead researcher. We would also like to remind you that you have the right to withdraw from this study at any time and without giving a reason.

What are the possible benefits of taking part?
The study will not have any direct benefit for you, but the information we get from this study may be helpful for the investigation into developing tests for people who have dementia but cannot speak English.

What will happen to the results of this study?
The lead researcher will write up the results of this study for her educational course, in the Doctorate in Clinical Psychology. It is also hoped that this research will also lead to publications. If you would like, we will give you a short report about how you performed on the tests. If you wish to receive a summary of the results please contact the lead researcher named at the end of this information.

Will any additional information be collected?
Information about your age, gender, general health, education level, occupation, ethnicity and marital status will need to be collected. This is important when conducting research, as people often want to know some information about the people that were involved in the research. If you are unable to provide this information a relative can do this on your behalf. We would need for you to consent for us to do this.

What if there is a problem?
If you have a concern about any aspect of this study, you can discuss this with the lead researcher or with her supervisors. If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital. In the event that something does go wrong and you are harmed during the research and this is due to someones negligence then you may have grounds for a legal action for compensation against Nottinghamshire Healthcare Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Who will have access to information about me?
Only the researchers involved in the running of this study will have access to your information. With your consent we will send a letter to your consultant confirming your participation in this study. This is standard practice for patients who are invited to take part in research studies.

Who is organising the research?
The study is being organised by the lead researcher and supervisors from the University of Nottingham.

Who has reviewed the study?
All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by a Research Ethics Committee in your local area.
In addition to this, research supervisors based at the University of Nottingham will monitor the research and provide supervision to the lead researcher. If you decide to participate you will be given a copy of this information sheet and a signed consent form to keep.

Contacts for further information

If you would prefer independent advice from a body that is not involved in this study, you can contact Alzheimer's Society or the Patient Advice and Liaison Services.

Lead Researcher Sobia Khan (Trainee Clinical Psychologist) University of Nottingham Mobile:	Clinical Research Supervisor Dr David Connelly: (Consultant Clinical Psychologist) Highbury Hospital **Tel:**
Academic Research Supervisor Dr Roshan Nair (Consultant Clinical Psychologist) University of Nottingham **Tel:**	
Independent agencies who you can contact for further advice about taking part in this study:	
Patient Advice and Liaison Services (PALS) PALS Duncan Macmillan House	Alzheimer's Society 7 Mansfield Rd Nottingham, NG1 3FB

Porchester Road, Nottingham NG3 6AA Freephone Helpline: 0800 0153367	0115 934 8468 **Tel:** 0115 934 8468

Thank you for taking the time to read this information sheet and considering taking part in this study.

Appendix 7: Capacity to consent assessment

The assessment of participant's capacity to consent

The assessment of dementia severity using non-verbal cognitive tests

What is the research about?

Why do you think you have been approached to take part in this research?

Do you have to be in this study if you do not want to participate?

What will you be asked to do if you take part in this research?

If you withdraw from this study will you still be able to receive regular treatment?

How much of your time do you think you would need to give to this research?

Who could you ask for more information about this research?

Who could you tell if you had concerns about taking part in this research?

Appendix 8: Consent form

CONSENT FORM

Title of Project: The assessment of dementia severity using non-verbal cognitive tests

Name of Lead Researcher: Sobia Khan (Trainee Clinical Psychologist)

		Please initial box
1	I confirm that I have read and understand the information sheet dated 09/09/09 (version 5) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	☐
3	I understand that sections of any of my research notes may be looked at by responsible individuals from Nottinghamshire Primary Care Trust or University of Nottingham where it is relevant to my taking part in research. I give permission for these individuals to have access to my records.	☐
4	I confirm that you can approach my carer for some information.	☐
5	I confirm that you can share information about my performance on the tests with a clinician who is involved in my care (e.g care co-ordinator, key worker, social worker, psychiatrist, nurse).	☐
6	I agree to take part in the above study	☐

Name of patient	Date	Signature
Researcher	Date	Signature

1 for patient; 1 for researcher; 1 to be kept with clinical notes

Appendix 9: Procedure and instructions for administering cognitive tests

Testing Procedure and Instructions

1. Go through participant information sheet
2. Capacity to consent assessment
3. Sign consent form
4. Building rapport

Remind participants that they can take a break at any point during testing. Make a note of when break is taken in data file.

Administered in private to ensure confidentiality, minimise distractions and reduce potential of discomfort.

Order of tests

1) Rey Complex Figure Test
2) Demographics (Demographics and MMSE to be administered to fill 30 minute gap until delayed trial on RCFT is to be administered)
3) Mini-Mental State Examination
4) Ravens Colour Progressive Matrices
5) Symbol Digit Modalities Test
6) Brixton Test
7) Clock Drawing Test
8) Colour Trails Test

Rey Complex Figure Test

Copy Trial: Look at this figure (point to the blank response sheet) I would like you to copy that figure onto this sheet of paper (point back to the stimulus card and say) copy it so that I know that this is the figure you drew. Do a good job. (After the participant has indicated that they have understood the instructions to the task, begin timing the drawing as the participant starts copying the figure.

Note time taken and day if testing – do not tell them you will be asking them to recall from memory

Note time when this task finished

Immediate recall: 3 minutes after – A short time ago I had you copy a figure. I would like you to draw that figure again, but this time from memory. (point to the blank response sheet and say) Draw that figure here.

Note time taken and day if testing – do not tell them you will be asking them to recall from memory

Delayed recall: 30 minutes after copy trial – A short time ago, I had you copy a figure. I would like you to draw that figure again, but this time from memory. (Point to the blank sheet and say) Draw that figure here.

Record the time in the day the trial commenced

Recognition Trial: Administered immediately after the Delayed Recall trial with no intervening task.

Some of the designs that are printed on these pages were part of the larger figure that I asked you to copy earlier. Circle the figure that were part of the larger design you copied. Each figure on these pages is facing the same direction as in the original, complete design. There are four pages, and the designs are numbered 1 to 24. Go ahead and begin.

Mini-Mental State Examination

Begin by asking the participant if they experience any problems with their memory. Then ask may I ask you some questions about your memory? These answers are not scored, however, they serve to orient the participant to the nature of the examination. The instructions are based on standardised instructions provided

1. **Orientation to time:** What is the year?

   What is the season?

   What is the day of the week?

   What is the date?

2. **Orientation to place:** Where are we now?

   What is the country?

   What is the city/ town?

   What part of the city are we now?

   What is the name of the building we are in now?

   What is the room we are in now?

3. **Registration:**

   Listen carefully. I am going to say three words. You say them back after I stop. Ready? Here they are....Apple, Pause, Penny, Pause, Table, Pause. Now repeat those words back to me. ..... Now keep these words in mind, I am going to ask you to say them again in a few minutes.

4. **Attention and Calculation:**

   Now I'd like you to subtract 7 from 100. Then keep subtracting 7 from each answer until I tell you to stop. What 100 take away 7?

   If participant refuses to perform the serial 7s task, substitute it with the following:

   Spell WORLD forward, then backward

5. **Recall:**

   What were those three words I asked you to remember?

6. **Naming:**

   What is this? Point to pen

   What is this? Point to watch

7. **Repetition:**

   Now I am going to ask you to repeat what I say. Read? "No ifs, ands or buts" now you say that.

   Repeat 5 times but only score the first trial

8. **Comprehension:**

   Detach next page. Tear it in half. Use the upper half of the page for comprehension, reading and drawing subtests. Use the lower half for the reading and Drawing items.

   Listen carefully because I am going to ask you to do something. Take this paper in your right hand, pause, fold it, pause, and put it on the table.

9. **Reading:**

   Please read this and do what it says.

10. Writing:

Please write a sentence, if they do not respond, say write about the weather.

Demographics sheet until 30 minutes lapses for **RCFT**

11. Drawing:

Please copy this design.

Ravens Colour Progressive Matrices

Look at this (point to upper figure) you see it is a pattern with a piece cut out of it. Each of these pieces (point to each in turn) is the right shape to fit the space, but only one of them is the right pattern. No1 is the right shape but is not the right pattern. No2 is not the right pattern at all. No3 is quite wrong. No6 is nearly right but it is wrong here (point to the white piece). Only one is right. Point to the piece which is quite right.

Now point to the piece which came out of this pattern

Symbol Digit Modalities Test

Please look at the boxes at the top of this page. You can see that each box in the upper row has a little mark in it. Now look at the boxes in the row just underneath the marks. Each of the boxes under the marks has a number. Each of the marks in the top row is different, and under each mark in the bottom row is a different number.

Now look at the next line of boxes. Just under the top two rows. Notice that the boxes on the top have marks, but the boxes underneath are empty. You are to fill each empty box with the number that should go there according to the way they are paired in the key at the top of the page. For example, if you look at the first mark, then look up at the key, you will see that the number 1 goes in the first empty box. Now what number should you put in the second box? That's right, so write number 5 in that box. What number goes in the next box? That's right! Now for practice fill in the rest of the boxes until you reach the double line. Then stop!

Now when I say "go" write in the numbers just like you have been doing and keep going as fast as you can until I say stop. When you come to the end of the line, go quickly to the next line without stopping. Do not erase as it will waste time, just keep working on the next ones as fast as you can.

Brixton Test

There are many pages here which all have the same basic design on them. There are always ten positions, and one of them is always coloured blue (point to the one filled in blue). However, the coloured one moves around according to various patterns that come and go without warning. These numbers (point to numbers underneath the circles) are just here to refer to the position – there is nothing complicated or mathematical about this test.

Now as I turn the pages, your job is to pick up on the pattern as best you can. And point to where you think the blue one is going to be on the next page. It's not guess work, you can work it out. For instances, imagine the blue one was here (point to 6) and then when I turn the page it goes to 7, and then to 8 and then to 9 – you might reasonably expect it next to go to 10.

From time to time the pattern changes without warning and then it is your job to pick up on the pattern as best as you can. Do you understand?

Obviously the first time you have nothing to go on, so your first answer will have to be a guess – have a guess as to where the blue one will be next.

Clock Drawing Test

Clock drawing

Imagine this is the face of a clock. Put the numbers on the clock face.

When finished, say:

Now put the hands on the clock to indicate 10 past 11.

Colour Trails Test 1

Practice Trial

In this box are different coloured circles with numbers in them. When I say "begin" I want you to take this pencil and connect the circles by going from 1 to 2 to 3 to 4 to 5 to 6 until you reach the end.

I want you to connect the circles in the correct order as quickly as you can, without lifting the pencil from the paper. If you make a mistake, I will point it out. When I do, I want you to move the pencil back to the last correct circle and continue from there. The line that you draw must go through the circles and must do so in the correct order. Do you have any questions?

Okay lets practice. Put your pencil here where this hand tells you to start. When I say "begin" connect the circles in order as quickly as you can until you reach the circle next to the hand telling you to stop. Ready? Begin.

Test Trial

Now I have a sheet with several more numbers and circles. Connect the circles in order like you did just a moment ago. Again, work as quickly as you can, and do not lift the pencil from the paper as you go. Make sure that your lines touch the circles.

You will start here where the hand tells you to start and end where the hand tells you to stop. Ready? Begin.

Colour Trails Test 2

Practice Trial

In this box are different coloured circles with numbers in them. This time I want you to take the pencil and connect the circles in order by going from this colour 1 (point to pink 1) to this colour 2 (point to yellow 2), to this colour 3 (point to pink 3) to this colour 4 (point to yellow 4) and so on until you reach the last number next to the hand telling you to stop.

Notice that the colour changes each time you go to the next number. I want you to work as quickly as you can. Do not lift the pencil from the paper once you have started. If you make a mistake, i will point it out. When I do, I want you to move the pencil to the last correct

circle and continue from there. As before, the line you draw must go through the circles in the correct order. Do you have any questions?

Practice

Okay, lets practice

Put your pencil here next to the hand telling you to start. When i say begin, connect the circles in order as quickly as you can, changing from one colour to the next, until you reach the hand telling you to stop. Ready? Begin.

Test Trial

Now I have a sheet with several more numbers and coloured circles. Connect the circles like you did just a moment ago. Again work as quickly as you can.

You will start here, where the hand tells you to start and end where the hand tells you to stop. Ready? Begin.

Appendix 10: Demographic questionnaire

Demographic Questionnaire

<table>
<tr><td colspan="3">Title: The assessment of dementia severity using non-verbal cognitive test</td></tr>
<tr><td colspan="2">Lead Researcher: Sobia Khan
(Trainee Clinical Psychologist)</td><td>Participant Code Number:</td></tr>
<tr><td>Dob:</td><td rowspan="6">Education:</td><td>Grammer School ☐ Comprehensive ☐</td></tr>
<tr><td>Age:</td><td>College ☐ Higher education ☐</td></tr>
<tr><td>Gender:

Male ☐ Female ☐</td><td>University ☐</td></tr>
<tr><td></td><td>Work based training/ apprenticeship ☐</td></tr>
<tr><td>Occupation / previous occupation (please state):</td><td>Other ☐
Please state:</td></tr>
<tr><td>Diagnosis:</td><td>How many years of education do you have?</td></tr>
<tr><td>When diagnosed:</td><td></td><td>What age did you complete your education?</td></tr>
<tr><td colspan="2">Ethnicity:
White ☐
British ☐
Irish ☐
Other White ☐</td><td>What was the highest qualification you have achieved?

Living arrangements:</td></tr>
<tr><td colspan="2">Mixed ☐
White and Black Caribbean ☐
White and Black African ☐
White and Asian ☐
Other Mixed ☐</td><td>Relationships status:
Single ☐
Married ☐
Divorced ☐
Widowed ☐
Other ☐ Please state:</td></tr>
<tr><td colspan="2">Asian or Asian British ☐
Indian ☐
Pakistani ☐
Bangladeshi ☐
Other Asian ☐</td><td>Medication:</td></tr>
<tr><td colspan="2">Black or Black British ☐
Caribbean ☐
African ☐
Other Black ☐</td><td>Co-morbidities:</td></tr>
<tr><td colspan="2">Chinese or other ethnic group ☐
Chinese ☐
Other ethnic group ☐</td><td>History of head injury:
Date of injury:</td></tr>
</table>

Appendix 11: Sample of feedback report for participant

Trent Doctorate of Clinical Psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

Tel:
Date

Address

Dear

I am writing to thank you for the valuable contribution you have made to my research. I understand that you have a review with your care co-ordinator soon and you requested for the outcomes of the assessments to be communicated to *.

As you may recall, the research was about trying to find different forms of assessment for older people who were experiencing changes in their memory, attention or other areas of their mental processing. I carried out seven tests with you on Date. Each of these assessed different mental processes and as promised, I will provide some feedback below.

Observations

During our time together, you were pleasant and appropriate in your social interaction. You commented that you had enjoyed the assessment process as it was a challenge. You struck me as someone who had some insight and awareness about the difficulties you were experiencing and you were taking an active role in managing the care and treatment you were receiving from services. For instance, you were able to evaluate what assessments had been completed so far and were able to tell me that * would be repeating assessments with you in a year. You were able to tell me that this was to monitor if there had been any change, in the difficulties you are experiencing, over time.

You had no problems with remembering that I was coming to see you or recalling the purpose of my research. Subsequently, you were assessed as having the capacity to consent to take part in this research.

Global cognitive ability

The Mini-Mental State Examination (MMSE) was used as a general measure of your thinking processes. On this, you scored 29 out of 30, which is within the 'normal range'. Although this is one point lower than when you were assessed previously on this test, it does not necessarily mean that you have experienced a decline in your thinking ability, as there can be some variation in scores across different assessors. A one point difference is not significant.

The Clock Drawing Test (CDT) was also administered. You completed this successfully without making any errors. This places you within the 'normal' range. The symbol Digit Modalities Test was also used as a screening measure of difficulties in thinking ability. On this you scored 31, this is again within the 'Normal range'.

As we discussed during our time together, people who are *highly educated* can achieve 30 out of 30 on the MMSE and make no errors on the CDT, but still have difficulties with their thinking ability. As you have 23 years of education, this would place you in the 'high level of education range'. It is worth taking this into consideration when interpreting any results.

Visuospatial Constructional ability and Visual Memory

To assess how you process visual information and your visual memory, I administered the Rey Complex Figure Test. On this test you scored within the 'Normal range' on visual-constructional ability. During our time together, you highlighted that your professional expertise lie in the field of Architecture. Therefore, we would expect someone with your experience to be highly skilled in drawing complex figures. When assessing how long it took you to copy the figure, your score again fell within the 'Normal' range. This suggests that your speed of information processing ability was adequate.

For immediate and delayed visual memory, you scored within the 'Above average range'. For recognition memory, you scored within the 'Average'. These scores suggest that your visual memory processing is within the 'Normal range'.

Attention and speed of information processing

The Colour Trails Test, part 1 and 2 were administered to assess your abilities in sustained visual attention, visual scanning and speed of information processing. Your scores were compared with a population of a similar age to you and with similar levels of education to you to interpret the score. On both of these tests you scored within the 'Below Average' range.

Executive processes

The Brixton test was administered to assess your executive processes. These are thinking processes which are related to every-day functioning ability. Some examples are abstract reasoning ability, making decisions, showing good judgment, maintaining attention, appropriate social behaviour, devising and following plans.

The Brixton test is a rule attainment test which can indicate if someone may have problems with executive processes that require further exploration. On this test you scored within the 'Moderate average range'.

Intellectual and reasoning ability

The Ravens Colour Progressive Matrices were administered to assess intellectual ability and the ability to think clearly. On this measure you scored 35 out of 36, this suggests that on this measure you scored within the 'Intellectually superior' range.

Summary

From the tests that were completed you presented with strengths in the area of visual-constructional ability, visual immediate, delayed and recognition memory and in reasoning ability. The scores for speed of information processing, attention and executive process were slightly weaker.

Conclusion

It is important to remember that the above assessments were completed as part of a research project. I collected the data in one visit with you and did not have access to a detailed history or any other relevant clinical information what is necessary to interpret test scores. In addition, I did not have any information to make comparisons about your abilities prior to the onset of your difficulties. Therefore, caution must be taken when interpreting the results.

I have enclosed a copy of the assessment report, which contains all the raw scores. I have also enclosed a copy the consent form which you signed on the day. I will forward a copy of these documents, together with a copy of this letter to * so that a copy can be placed within your clinical records.

I would like to take this opportunity to thank you once again for your support with this research.

Yours sincerely

Sobia Khan
Lead Researcher & Trainee Clinical Psychologist

Cc

Enc
Assessment Report
Consent form

Assessment Report

Name:
Address:
Date of birth:
Reason for referral: Research
Referred by:
Date assessed:

Rey Complex Figure Test

The purpose of this test is to assess visual-spatial constructional ability, speed of processing, visual immediate, delayed and recognition memory.

	Copy	**Time to copy**	**Immediate recall**	**Delayed recall**	**Recognition total correct**
Raw score	33	211	23.5	22.5	20
T score	-	-	78	76	50
Percentile	>16	>16	>99	>99	50
Suggested clinical interpretation	Normal range	Normal range on speed of processing	Above Average	Above Average	Average

Range of clinical interpretation

***T*-score range**	**Suggested clinical interpretation**
>55	Above average
45 to 54	Average
40 to 44	Below average
35 to 39	Mildly impaired
30 to 34	Mild- to- Moderately impaired
25 to 29	Moderately impaired
20 to 24	Moderately-to-severely impaired
<19	Severely impaired

Mini-Mental State Examination

This is a screening measure to assess for cognitive impairment particularly in older people.

Domain	Score
Orientation to time	5
Orientation to place	5
Registration	3
Attention and calculation	5
Recall	3
Naming	2
Repetition	0
Comprehension	3
Reading	1
Writing	1
Drawing	1
Total	29/30

Cut off	Classification
27-30	Normal
21-26	Mild
11-20	Moderate
0-10	Severe

Raven's Coloured Progressive Matrices

This test assesses reasoning ability within the visual modality and intellectual ability.

Score	Percentile	Grade
35/ 36	95 percentile	Intellectually Superior

Colour Trails Test

This test assesses speed of attention, sequencing, mental flexibility, visual search, and motor function.

	Raw score	Standard score	*T* score	Percentile score	Clinical Impression
Colour Trails 1	73 seconds	90	43	24	Below average
Colour Trails 2	143 seconds	83	39	14	Below Average

Brixton spatial anticipation test

This is a concept (or 'rule') attainment task. This task measures one aspect of executive functioning.

Raw score	Scaled score	Classification
19	5	Moderate Rang

Raw Score	Scaled Score	Classification
0-7	10	Very superior
8	9	Superior
9-10	8	Good
11-13	7	High Average
14-17	6	Average
18-20	5	Moderate Average
21-23	4	Low average
24-25	3	Poor
26-31	2	Abnormal
>31	1	Impaired

Symbol Digit Modalities Test

This is a measure of attention, visual scanning, concentration, and motor and psychomotor speed.

Total score	Difference from sample mean	SD (1)	Interpretation	Clinical Impression
31	-	No norms were available for age group. However, this score was equal to what was expected of someone between 65 -78 years. Mean score was 32.28 (SD 11.27).		

Score	Interpretation
Scores approximately 1 standard deviation (Number in brackets) below the mean for a particular age group at a particular education level.	Low scores
Scores approximately 1.5 standard deviations below the mean for a particular age group at a particular education level.	Moderately low scores
Scores approximately 2 standard deviations below the mean for a particular age group at a particular education level.	Very low scores

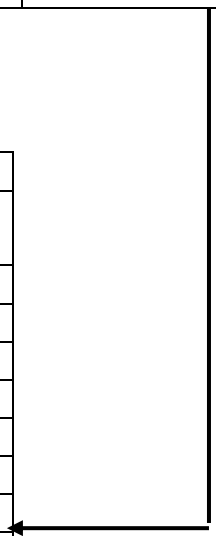
Clock Drawing Test

This tests assesses for dementia as well as for visual-spatial, constructional, and executive difficulties. Higher T score values indicate higher levels of cognitive impairment. The percentiles indicate what percentage of sample of older people would score more than what was achieved by participant.

Errors	***T* Score**	**Percentile**	**Clinical Impressions**
0	37	16	Very good performance

Range of *T* scores

Percentile equivalents for T-scores	
T Scores	Approximate percentiles
70	98
65	93
60	84
55	69
50	50
45	31
40	16
35	7
30	2



Appendix 12: Sample of feedback report for referrer

Trent Doctorate of Clinical Psychology
Institute of Work Health and Organisations
University of Nottingham
International House, B Floor
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

Tel:
25th September 2009

The cognitive assessment of dementia severity using non-verbal cognitive tests

Dear

Re:

Thank you for approaching your client and contributing to the above research. * provided consent for me to share her results on the cognitive tests with you. I have enclosed a copy of the consent form and I would appreciate it if you could place this within her clinical records.

The assessment report will provide a brief outline of the cognitive domains which each of the tests measures. I have also included the scores for each test and have documented the clinical impressions where possible.

Please do not hesitate to contact me on the number above should you want to discuss these results further.

Yours sincerely

Sobia Khan
Lead researcher & Trainee Clinical psychologist

Supervised by:

Dr Roshan Das Nair
Dr Helen Philpott

Assessment Report

Name:
Address:
Date of birth:
Reason for referral: Assessment of dementia severity research study
Referred by:
Date assessed:

Observations:

Rey Complex Figure Test

The purpose of this test is to assess visual-spatial constructional ability, speed of information processing and visual memory.

	Copy	**Time to copy**	**Immediate recall**	**Delayed recall**	**Recognition total correct**
Raw score	32	404 Seconds	2.5	0	15
T score	-	-	21	<20	<20
Percentile	>16	11-16	<1	<1	<1
Suggested clinical interpretation	Normal range - Intact visuospatial constructional skills	Slower speed of information processing	Moderately-to-severely impaired	Severely impaired	Severely impaired

Range of clinical interpretation

***T*-score range**	**Suggested clinical interpretation**
>55	Above average
45 to 54	Average
40 to 44	Below average
35 to 39	Mildly impaired
30 to 34	Mild- to- Moderately impaired
25 to 29	Moderately impaired
20 to 24	Moderately-to-severely impaired
<19	Severely impaired

Mini-Mental State Examination

This is a screening measure to assess for cognitive impairment particularly in older people.

Domain	Score
Orientation to time	4
Orientation to place	5
Registration	3
Attention and calculation	4
Recall	1
Naming	2
Repetition	0
Comprehension	3
Reading	1
Writing	1
Drawing	1
Total	25/30

Cut off	Classification
27-30	Normal
21-26	Mild
11-20	Moderate
0-10	Severe

Raven's Coloured Progressive Matrices

This test assesses reasoning ability within the visual modality.

Score	Percentile	Grade
11/ 36	5th percentile	Intellectually impaired range

Colour Trails Test

This test assesses speed of attention, sequencing, mental flexibility, visual search, and motor function.

	Raw score	Standard score	*T* score	Percentile score	Clinical Impression
Colour Trails 1	78 seconds	90	43	24	Below average
Colour Trails 2	200 seconds	78	35	7	Impaired range

Brixton spatial anticipation test

This is a concept (or 'rule') attainment task. This task measures one aspect of executive functioning.

Raw score	Scaled score	Classification
21	4	Low average

Symbol Digit Modalities Test

This test is used to assess divided attention, visual scanning, tracking, and motor speed.

Total score	Difference from sample mean	SD (1)	Interpretation	Clinical Impression
13	32.28	-2.8 (11.27)	Very low scores	Suggestive of cognitive difficulties

Score	Interpretation
Scores approximately 1 standard deviation (Number in brackets) below the mean for a particular age group at a particular education level.	Low scores
Scores approximately 1.5 standard deviations below the mean for a particular age group at a particular education level.	Moderately low scores
Scores approximately 2 standard deviations below the mean for a particular age group at a particular education level.	Very low scores

Clock Drawing Test

This tests assesses for dementia as well as for visual-spatial, constructional, and executive difficulties. Higher T score values indicate higher levels of cognitive impairment.

Errors	***T* Score**	**Percentile**	**Clinical Impressions**
1	49	50	Quite good performance

Range of *T* scores

Percentile equivalents for T-scores	
T Scores	Approximate percentiles
70	98
65	93
60	84
55	69
50	50
45	31
40	16
35	7
30	2

Appendix 13: Flow chart of recruitment