# REVIEW

# Neuropsychological deficits in frontotemporal dementia and Alzheimer's disease: a meta-analytic review

## A D Hutchinson, J L Mathias

We sought to identify the cognitive tests that best discriminate between Alzheimer's disease (AD) and frontotemporal dementia (FTD). A comprehensive search of all studies examining the cognitive performance of persons diagnosed with AD and FTD, published between 1980 and 2006, was conducted. Ninety-four studies were identified, comprising 2936 AD participants and 1748 FTD participants. Weighted Cohen's *d* effect sizes, percentage overlap statistics, confidence intervals and fail-safe *N*s were calculated for each cognitive test that was used by two or more studies. The most discriminating cognitive tests were measures of orientation, memory, language, visuomotor function and general cognitive ability. Although there were large and significant differences between groups on these measures, there was substantial overlap in the scores of the AD and FTD groups. Age, education, years since diagnosis and diagnostic criteria did not significantly contribute to the group differences. Given the large overlap in the test performance of persons diagnosed with AD and FTD, cognitive tests should be used cautiously and in conjunction with a medical history, behavioural observations, imaging and information from relatives when making differential diagnoses.

.......................................................................................

D
ementia is an important social and health issue in our ageing population, with the prevalence of Alzheimer's disease (AD) in the USA predicted to increase by approximately 300% by 2050.[1] Two forms of dementia, AD and frontotemporal dementia (FTD), and their differential diagnosis using cognitive tests, are addressed in the current study.

AD is the most prevalent and well researched type of dementia.[2] People with AD usually exhibit a progressive decline in general cognitive functioning,[3] together with acalculia, apraxia, visuospatial problems, poor orientation and impaired language comprehension, attention and memory.[2 4–6] Memory problems are one of the earliest symptoms, with the ability to learn and retrieve information being affected initially and impairments in recognition memory developing later.[2 3 7]

FTD, on the other hand, is variously referred to as Pick disease, dementia of the frontal type, frontal lobe dementia of the non-Alzheimer's type, frontotemporal lobar degeneration, semantic dementia and progressive non-fluent aphasia.[8 9] FTD refers to a group of degenerative dementias that are characterised by atrophy of the fronto-temporal cortex.[5 8 10] Although people with FTD

may also experience memory problems, these deficits are more frequently associated with a tendency to respond impulsively and a failure to monitor performance.[5 9] As with AD, retrieval of information from memory is affected. However, unlike AD, the provision of cues can assist performance.[9] Persons with FTD also have more impairments in executive functioning[2–6 9 11 12] and language (eg, inability to repeat, stereotyped phrases, reduced speech).[5 7 9] Finally, although general cognitive ability may be unaffected in the early stages of FTD, it does decline over time.[4 12]

While definitive diagnoses of AD and FTD can only be made at autopsy, interim diagnoses usually take into account the base rates of the disorder, clinical criteria, medical history, physical examination, brain imaging and neuropsychological assessments.[6 10 12] Unfortunately, the diagnosis of FTD can be difficult because of its insidious and gradual onset[8 9] and can also be misdiagnosed as AD.[13 14] However, accurate differential diagnosis has become increasingly important because of the recent availability of pharmacological treatments that temporarily improve the cognitive and functional abilities of people with AD.[15–21] Moreover, delays in the commencement of these treatments significantly reduce their benefits,[19 22] highlighting the importance of early and accurate diagnosis. In contrast, no specific pharmacological treatments currently exist for FTD, although research is continuing.[23]

If neuropsychological assessments are to make a valuable contribution to the diagnosis of dementia, clinicians need to know what cognitive tests best discriminate between the different types of dementia. Despite the differing clinical characteristics of these two disorders and the wide variety of tests that have been researched with AD and FTD samples, research has yielded inconsistent findings, both in terms of the cognitive constructs (eg, memory, language) and the specific tests that appear to differentiate between AD and FTD.[14 24–33] These inconsistencies may, in part, be due to methodological differences in the research literature, such as variations in the participants' age, time since diagnosis, stage of dementia and diagnostic criteria. A recent narrative review of the literature concluded that there are no tests that

See end of article for authors' affiliations
......................

Correspondence to:
Dr J Mathias, School of Psychology, University of Adelaide, Adelaide, SA, 5005, Australia; jane.mathias@adelaide.edu.au

........................................................

**Abbreviations:** ACE, Addenbrooke's Cognitive Examination; AD, Alzheimer's disease; AVLT, Auditory Verbal Learning Test; FTD, frontotemporal dementia; MMSE, Mini-Mental State Examination; Nfs, fail safe N; %OL, percentage overlap; WAB, Western Aphasia Battery; WLS, weighted least squares; WMS, Wechsler Memory Scale

accurately discriminate between AD and FTD.[34] However, a review of this type cannot adequately consolidate and analyse the data or reconcile differences in methodology.[35] A meta-analysis, in contrast, provides an objective and quantitative means by which to directly compare existing findings, thereby providing an important addition to the existing literature.[36][37] The current study therefore undertook a meta-analytic review of research that has compared the cognitive performance of AD and FTD samples in order to identify the tests that best discriminate between AD and FTD and, therefore, are most likely to be of use when making a differential diagnosis.

## METHOD
### Literature search and inclusion criteria
A comprehensive search of the PubMed and PsycINFO electronic databases from January 1980 to November 2006 was undertaken to identify all published studies that assessed the cognitive functioning of AD and FTD samples. The key search terms that were used to capture relevant articles are shown in table 1. The bibliographies of relevant papers were also examined for additional references. To be selected for the current meta-analysis, a study had to meet the following inclusion criteria: (1) it examined groups of participants with AD and FTD (excludes case studies), (2) cognitive tests were administered to both groups (excludes rating scales and questionnaires), (3) the cognitive tests were not used for the diagnosis and classification of participants into the AD and FTD groups (ie, a test could not be used as both a dependent and independent variable), (4) statistical data enabling the calculation of Cohen's *d* effect sizes[38] was provided (eg, mean and SD, t tests or one way F tests) and (5) it was published in English.

This literature search yielded 2785 potentially relevant studies, 93 of which met all of the inclusion criteria. These 93 studies used a total of 136 different cognitive tests which, in turn, yielded results for 1019 different test scores (ie, some tests yielded multiple scores). A total of 114 of these test scores were used by more than one study and were, therefore, included in the current meta-analysis. Of the studies that did not meet one or more of the inclusion criteria, 2687 did not undertake cognitive testing with the AD and FTD samples, 67 did not provide data that would enable the calculation of effect sizes and 96 studies only used cognitive tests for the diagnosis and classification of participants. An additional paper[39] was excluded because it presented data that had been reported in another publication that was to be included in this study.[40] This

was done to ensure that the data were independent, as meta-analytic techniques assume that a single study only contributes one score to the calculation of a particular effect size.[41] Finally, one publication was treated as two studies because data were provided for separate samples in the one article.[42] Thus 94 independent studies were effectively included in the final meta-analysis.

### Data collection and preparation
Consistent with the most commonly used diagnostic criteria for AD,[43] the Mini-Mental State Examination (MMSE) was frequently used when diagnosing dementia. The MMSE was therefore only included as a dependent variable in the current meta-analysis when groups were formed on the basis of a post-mortem diagnosis of AD or FTD dementia. However, where provided, MMSE scores are reported as descriptive data for the AD and FTD samples (refer to table 2). In addition, whenever it was not clear whether a particular cognitive test was used to diagnose dementia and classify participants, this test was excluded from the meta-analysis in order to ensure that a measure was not used as both a dependent and independent variable.

Where test scores were obtained from different editions of the same test (eg, Wechsler Adult Intelligence Scales), these were combined for the purposes of calculating mean effect sizes. Care was taken to ensure that the scores for a given test provided independent measures of performance. Thus sub-scores and total scores from a test could not both be used in the calculation of an effect size. Similarly, where a study provided multiple scores for the same test (eg, easy and hard versions of a geometric figure task[44]), effect sizes were calculated for each score and then averaged to provide a single effect size for that study; thereby ensuring that a study only contributed one score to the calculation of a mean effect size. This was not necessary when effect sizes for different scores from the same test were reported separately (eg, speed and accuracy).

All tests were broadly grouped into seven cognitive categories in order to organise the findings of this meta-analysis. Lezak *et al*[45] and Spreen and Strauss[46] were consulted for this purpose. These categories are: orientation and attention, memory, construction, verbal abilities and language, concept formation and reasoning, executive functioning and motor tasks, and general ability and intelligence.

### Effect size calculations and analyses
Cohen's *d*[38] effect sizes were calculated to measure the extent of the difference between the scores of the AD and FTD samples, where a small effect size is defined as *d* = 0.2, a medium effect as *d* = 0.5 and a large effect as *d* = 0.8. To put this into perspective, an effect size of 0.5 indicates that the scores for the two groups differ by half of a pooled standard deviation. The percentage overlap between the scores of the two groups can also be calculated from an effect size[47]; *d* = 0 equates to 100% overlap (ie, the groups are indistinguishable), *d* = 1.0 equates to 45% overlap and *d* = 3.0 equates to less than 5% overlap in the groups' scores. These statistics are provided for all effect sizes in order to facilitate their interpretation, especially with regard to their potential for assisting with differential diagnosis.

Effect sizes were calculated in a multi-stage process. The first stage involved calculating effect sizes for each score of every test that was used by each individual study. When calculating effect sizes, FTD scores were always subtracted from AD scores. In most cases, higher test scores indicated better performance. Therefore, a positive effect size indicates that FTD participants were more impaired on a given measure than AD participants. In cases where a higher score indicated greater impairment than a lower score (eg, errors, speed), the direction of the effect sizes for these scores was transformed so that a positive effect size still indicated greater impairment in the FTD group.

---

**Table 1** Key search terms

| Alzheimer's disease | Frontotemporal dementia | Assessment |
|---|---|---|
| Alzheimer's disease | Frontotemporal dementia | Diagnosis |
| Alzheimer* | Frontal dementia | Screening |
|  | Pick's disease | Cognit* |
| DAT | Semantic dementia | Neuropsychol* |
| AD | Progressive nonfluent aphasia |  |
|  | Primary progressive aphasia |  |
|  | FTLD |  |
|  | FLD |  |
|  | FTD |  |
|  | DFT |  |
|  | PNFA |  |
|  | PPA |  |

AD, Alzheimer's disease; DAT, dementia of the Alzheimer's type; DFT, dementia of the frontal type; FLD, frontal lobe dementia of the non-Alzheimer's type; FTD, frontotemporal dementia; FTLD, frontotemporal lobe dementia; PNFA, progressive non-fluent aphasia; PPA, primary progressive aphasia.
*Search term and its derivatives were used (eg, cognition, cognitive, neuropsychology, neuropsychological).

**Table 2** Demographic details for the Alzheimer's disease and frontotemporal dementia groups

| | **AD** | | | | | | **FTD** | | | | |
| | **N* studies** | **N* participants** | **Mean** | **SD** | **Range** | | **N* studies** | **N* participants** | **Mean** | **SD** | **Range** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 94 | 2936 | 31.2 | 41.1 | 4.0 | 236.0 | 94 | 1748 | 18.6 | 12.6 | 2.0 | 67.0 |
| Age (y) | 88 | 2792 | 70.6 | 5.1 | 55.7 | 79.9 | 89 | 1658 | 64.8 | 4.6 | 56.0 | 78.1 |
| Education (y) | 61 | 1878 | 11.4 | 3.0 | 4.0 | 16.3 | 62 | 1091 | 11.8 | 3.0 | 4.0 | 17.4 |
| Time since diagnosis (y) | 38 | 1237 | 3.8 | 1.5 | 1.7 | 9.8 | 38 | 752 | 3.9 | 1.4 | 1.3 | 8.4 |
| MMSE | 76 | 2506 | 21.4 | 3.0 | 8.9 | 26.6 | 75 | 1460 | 22.7 | 3.0 | 13.8 | 28.7 |
| CDR | 20 | 398 | 1.2 | 0.7 | 0.5 | 3.5 | 19 | 332 | 1.2 | 0.9 | 0.5 | 4.5 |

AD, Alzheimer's disease; CDR, Clinical Dementia Rating scale; FTD, frontotemporal dementia; MMSE, Mini-Mental State Examination.
*The N differs between variables because not all studies provided data for each variable

The effect sizes for all studies that used a particular test score were aggregated to calculate a mean effect size (and SD). Whereas mean effect sizes measure the extent (and direction) of the difference between AD and FTD, the SD indicates the variation in effect sizes between studies. Before aggregating the scores from individual studies, each effect size was weighted to take into account the fact that the reliability of an effect size is affected by the size of the sample from which it is derived. According to Lipsey and Wilson,[48] the inverse of the variance provides a better measure of the precision of an effect size than sample size. Consequently, the inverse of the variance was used to weight all effect sizes ($d_w$) using the formulae below.

$$d_w = \frac{\sum wES}{\sum w}$$

where ES = effect size,

$$w = \frac{1}{SE^2}$$

and $SE$ = standard error.

Effect sizes are only reported for tests that were used by two or more studies because effect sizes that are based on a single study are not thought to provide a reliable measure of group differences.[41] Ninety-five per cent confidence intervals (CIs) for these mean effect sizes were additionally calculated in order to determine their statistical significance. If a CI does not span zero, this indicates that the true population effect size differs significantly from zero, indicating that there is a significant difference between the performance of the AD and FTD groups.

One problem facing all meta-analyses is that studies with statistically significant results are more likely to be published and, therefore, more likely to be included in a meta-analysis. Failure to include unpublished studies with non-significant results increases the risk of a type 1 error, which may result in an effect size being overestimated.[49] A fail safe N (Nfs) was therefore calculated, using the method described by Rosenthal,[41] to address this problem. This statistic estimates the number of unpublished studies, with non-significant results (ie, small effect) that would be required in order to call the current findings into question.[41] The larger the number, the more confident we can be in a particular finding.

An additional problem that faces meta-analytic studies (and qualitative reviews) is that there are methodological differences between studies that may affect the findings and, therefore, the effect sizes. The impact of diagnostic certainty was therefore examined in order to determine whether the research findings were affected by the criteria that were used to establish

**Table 3** Orientation and attention: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | **N studies** | **N participants** | **Mean Cohen's $d_w$** | **SD $d_w$ (95% CI)** | **Min $d$** | **Max $d$** | **Nfs** | **%OL** | **Study references** |
|---|---|---|---|---|---|---|---|---|---|
| ACE Orientation | 2 | 170 | −1.08 | 0.25 (−1.43, −0.73) | −1.20 | −0.95 | 9 | 41 | 85 102 |
| ACE Attention | 2 | 170 | −0.79 | 0.24 (−1.13, −0.45) | −0.86 | −0.71 | 6 | 53 | 85 102 |
| Trails B (errors) | 2 | 62 | −0.60 | 0.37 (−1.11, −0.09) | −0.75 | −0.20 | 4 | 62 | 74 123 |
| Stroop (word reading/Stroop A) | 4 | 105 | 0.46 | 0.45 (0.02, 0.89) | −0.40 | 5.2 | 5 | 67 | 31 60 86 131 |
| Digit Symbol | 5 | 373 | −0.40 | 0.33 (−0.69, −0.11) | −0.67 | −0.28 | 5 | 73 | 56 62 64 86 94 |
| Modified Trails (connection/min) | 4 | 225 | −0.38 | 0.28 (−0.66, −0.10) | −0.69 | −0.18 | 4 | 73 | 25 93 97 98 |
| Stroop (number correct—incongruent trial) | 6 | 358 | −0.29 | 0.29 (−0.52, −0.06) | −0.57 | 0.41 | 3 | 79 | 28 44 60 93 97 98 |
| Digit Span (overall score) | 8 | 479 | 0.29 | 0.34 (0.05, 0.52) | −0.47 | 0.98 | 3 | 79 | 27 29 62 68 86 126 130 134 |
| Corsi Blocks (visual span) | 4 | 166 | −0.29 | 0.34 (−0.62, 0.04) | −0.73 | −0.06 | 2 | 79 | 29 68 129 130 |
| Trails B (time) | 8 | 233 | −0.28 | 0.39 (−0.56, −0.01) | −0.89 | 2.39 | 3 | 79 | 60 74 80 108 111 114 123 129 |
| Attentional Matrices | 3 | 99 | −0.22 | 0.35 (−0.62, 0.18) | −0.45 | 0.27 | 0 | 85 | 68 123 130 |
| Stroop (colour naming/Stroop B) | 3 | 92 | 0.22 | 0.40 (−0.22, 0.67) | −0.39 | 1.53 | 0 | 85 | 31 60 86 |
| Stroop (interference) | 2 | 70 | 0.18 | 0.36 (−0.33, 0.68) | −0.07 | 0.49 | 0 | 85 | 88 100 |
| Stroop (errors—incongruent trial) | 3 | 78 | 0.18 | 0.40 (−0.28, 0.63) | −0.13 | 1.09 | 0 | 85 | 27 80 130 |
| Digit Span (forwards) | 13 | 453 | 0.14 | 0.36 (−0.06, 0.33) | −0.80 | 2.73 | 0 | 92 | 28 30 60 72 74 78 88 91 110 111 119 123 129 |
| Stroop (time—incongruent trial) | 2 | 40 | 0.14 | 0.45 (−0.48, 0.77) | −0.23 | 0.37 | 0 | 92 | 27 130 |
| Trails A (time) | 5 | 162 | −0.12 | 0.38 (−0.46, 0.21) | −0.73 | 0.32 | 0 | 92 | 60 74 86 111 129 |
| Digit Span (backwards) | 19 | 697 | 0.11 | 0.35 (−0.04, 0.27) | −1.16 | 1.08 | 0 | 92 | 25 28 30 60 67 72 74 77 78 80 88 91 93 97 98 110 111 119 123 |

ACE, Addenbrooke's Cognitive Examination; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

**Table 4** Perception: weighted Cohen's d effect sizes* for each test, rank ordered by size

| | N studies | N participants | Mean Cohen's $d_w$ | SD $d_w$(95% CI) | Min d | Max d | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| VOSP position discrimination | 2 | 57 | −0.70 | 0.40 (−1.25, −0.14) | −1.34 | 0.00 | 5 | 57 | 28 109 |
| VOSP dot counting | 3 | 78 | −0.68 | 0.54 (−1.29, −0.06) | −1.07 | 0.00 | 7 | 57 | 109 110 120 |
| VOSP object decision | 4 | 109 | 0.34 | 0.40 (−0.06, 0.73) | −0.75 | 0.99 | 3 | 79 | 78 109 110 120 |
| VOSP cube analysis | 4 | 98 | −0.23 | 0.42 (−0.64, 0.17) | −0.61 | 0.52 | 1 | 85 | 109 110 119 120 |
| VOSP incomplete letters | 4 | 108 | 0.04 | 0.39 (−0.35, 0.43) | −0.83 | 0.47 | 0 | 100 | 78 109 110 119 |

Max d, maximum effect size; Mean $d_w$, mean weighted effect size; Min d, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; SD $d_w$, SD of weighted effect size; %OL, per cent overlap between Alzheimer's disease and frontotemporal dementia groups; VOSP, Visual Object and Space Perception Battery; 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

dementia type. Every study was given a score according to the diagnostic criteria that were used to define each of their samples. As post-mortem neuropathological confirmation of AD or FTD provides the gold standard for a diagnosis of dementia, this was given a score of three ($N_{AD}$ = 5, $N_{FTD}$ = 5 studies), while studies using the National Institute of Neurologic, Communicative Disorders and Stroke-AD and Related Disorders Association (NINCDS-ADRDA) diagnostic criteria for AD[43] and the Lund and Manchester criteria for FTD[50] were given a score of 2 ($N_{AD}$ = 82, $N_{FTD}$ = 74 studies). Studies that used revised or similar criteria, particularly prior to the publication of these guidelines, were also given a score of 2. Finally, a score of 1 was given to studies whose diagnostic criteria were less rigorous or not clearly specified ($N_{AD}$ = 7, $N_{FTD}$ = 15 studies). The diagnostic scores for studies therefore varied between 2 and 6.

The potential influence of three other methodological variables (participant age, years since diagnosis and education) was also examined. Mean age, years since diagnosis and years of education were calculated for each study (eg, $M_{AD+FTDage}$) for this purpose. This was done by combining the data from the AD and FTD groups for that study and weighting it by the sample sizes of the AD and FTD groups.

While there is considerable confusion within the literature regarding the best way to analyse the influence of methodological variables on effect sizes, Steel and Kammeyer-Mueller[51] have demonstrated that a weighted least squares (WLS) multiple regression (using the inverse of the sampling error variance to weight the multiple regression) is less affected by the problems of heteroscedasticity and multicollinearity, which can affect the accuracy of these analyses. Heteroscedasticity occurs when the distribution of study sample sizes is skewed (eg, a large number of studies with small samples and few with very large samples) and multicollinearity occurs when predictor variables are correlated with one another.[51 52] A WLS multiple regression was therefore conducted, in addition to Pearson r correlation coefficients, in order to examine these methodological variables. Standard errors and confidence intervals for the results of the WLS multiple regression were adjusted according to the method outlined by Hedges.[53]

## RESULTS
### Participants
The demographic characteristics for the participants in the 94 studies that were included in the current meta-analysis are shown in table 2. Overall, 4684 participants were included in this study. Gender was only reported in 69 studies, providing data for 3480 cases (males: $N_{AD}$ = 917, $N_{FTD}$ = 737; females: $N_{AD}$ = 1253, $N_{FTD}$ = 573). When the demographic characteristics of the AD and FTD participants that were included in this meta-analysis were compared, the FTD participants were found to be significantly younger (t = 12.59, df = 86, p<0.001), more educated (t = −2.51, df = 60, p = 0.015) and had significantly higher

MMSE scores (t = −5.41, df = 74, p<0.001) than the participants with AD. However, the two groups did not differ significantly in terms of the time since their diagnosis (t = 0.38, df = 36, p = 0.709) or Clinical Dementia Rating scores (t = 0.25, df = 18, p = 0.807).

Five studies had pathological confirmation of AD,[54–58] 82 used published criteria,[14 25 27–31 40 42 44 59–129] and seven did not specify diagnostic criteria.[130–136] For the diagnosis of FTD, five studies had pathological confirmation,[54–58] 74 used published criteria,[14 25 27 28 30 31 40 42 44 59–64 66–73 76–105 108–116 118 120–124 126–129 131] 14 did not clearly specify the diagnostic criteria[29 65 74 75 106 107 117 125 130 132–136] and one study used a combination of clinical observations and published criteria.[119]

There were not enough studies reporting data for premorbid intelligence (n = 6) or scores on the Dementia Rating Scale (n = 6) to report or analyse this information. Similarly, too few studies reported information about participants' current medications (n = 19), much of which was not comparable, to meaningfully examine this information. Thirty-five studies reported information regarding the inclusion or exclusion of depressed participants. However, only six of these provided scores from depression rating scales.

### Cognitive tests
The weighted effect sizes ($d_w$) for all measures (mean, SD, 95% CI, minimum, maximum), grouped according to test category (orientation and attention, memory, construction, verbal abilities and language, concept formation and reasoning, executive functioning and motor tasks, and general ability and intelligence) and rank ordered by size, are provided in tables 3–10. Nfs and the percentage overlap between groups (%OL) are also provided, as are the number of studies (N) that used each measure, the number of participants that were assessed and the study references. The conclusions drawn from this study are based on the combined interpretation of these statistics. From a neuropsychological perspective, differential diagnoses are likely to be more accurate when there are large group differences (d) and there is limited overlap in the cognitive profiles (%OL). A clinician would also be more confident in their decision if the CI did not span zero (ie, groups differ significantly) and it is unlikely that the possibility of unpublished findings would draw a diagnosis into question (Nfs). Thus in order for a measure to be useful for differentiating between AD and FTD dementia, it had to meet the following criteria: (i) have a large effect size (ie, $d_w \geqslant 0.8$) and, consequently, a limited degree of group overlap (%OL <50), (ii) a confidence interval that did not span zero (CIs are affected by N and SD/range) and (iii) an Nfs score that was large enough to make it unlikely that there were that number of unpublished studies with non-significant findings in existence. As different tests were used with varying frequency, it was decided that the Nfs should be greater than the number of published studies that had used the test (ie, Nfs >$N_{studies}$).

**Table 5** Memory: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N studies | N participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min *d* | Max *d* | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| AVLT (recognition) | 2 | 46 | −1.25 | 0.59 (−2.08, −0.43) | −4.0 | 0.00 | 11 | 35 | 27 134 |
| Rey Complex Figure (recall) | 21 | 783 | −1.09 | 0.37 (−1.25, −0.93) | −6.51 | 0.07 | 93 | 41 | 25, 27, 29–31, 68, 78, 82, 86, 89, 90, 93, 97, 98, 110, 115, 119, 120, 123, 129, 134 |
| Recognition Memory Test (words) | 4 | 113 | −1.03 | 0.41 (−1.43, −0.63) | −1.46 | −0.55 | 17 | 45 | 28 82 119 120 |
| WMS Logical Memory (% retention) | 3 | 84 | −1.00 | 0.43 (−1.48, −0.51) | −3.21 | −0.21 | 12 | 45 | 27 28 31 |
| ACE Memory | 2 | 170 | −0.92 | 0.25 (−1.27, −0.58) | −1.05 | −0.78 | 7 | 48 | 85 102 |
| WMS Logical Memory (delayed) | 10 | 303 | −0.90 | 0.42 (−1.16, −0.65) | −2.17 | 0.30 | 35 | 48 | 30 42 56 78 82 86 94 110 124 |
| WMS Visual Reproduction (delayed) | 3 | 95 | −0.90 | 0.41 (−1.37, −0.44) | −1.23 | 0.48 | 11 | 48 | 56 94 124 |
| Selective Reminding Test (total recall) | 2 | 46 | −0.89 | 0.46 (−1.53, −0.25) | −1.06 | −0.68 | 7 | 48 | 60 121 |
| AVLT (delayed recall) | 5 | 235 | −0.87 | 0.39 (−1.21, −0.53) | −4.32 | −0.13 | 17 | 48 | 27 96 101 123 134 |
| AVLT (immediate recall) | 5 | 235 | −0.76 | 0.37 (−1.09, −0.43) | −2.26 | −0.04 | 14 | 53 | 27 96 101 123 134 |
| CVLT (delayed free recall) | 5 | 224 | −0.68 | 0.32 (−0.96, −0.40) | −1.39 | −0.49 | 12 | 57 | 87 93 97 98 111 |
| WMS Logical Memory (immediate recall) | 11 | 335 | −0.63 | 0.40 (−0.86, −0.40) | −2.23 | 0.25 | 24 | 62 | 30, 31, 42, 56, 78, 82, 94, 110, 120, 124 |
| Babcock Story Recall Test | 2 | 74 | −0.58 | 0.34 (−1.04, −0.11) | −0.81 | −0.23 | 4 | 62 | 68 123 |
| Selective Reminding Test (recognition) | 2 | 46 | −0.52 | 0.45 (−1.13, 0.10) | −0.59 | −0.42 | 3 | 67 | 60 121 |
| CERAD−NP (delayed recall) | 2 | 193 | −0.50 | 0.21 (−0.79, −0.21) | −0.85 | −0.35 | 3 | 67 | 70 105 |
| WMS Visual Reproduction (immediate) | 3 | 101 | −0.48 | 0.38 (−0.91, −0.05) | −0.80 | −0.08 | 4 | 67 | 56 94 124 |
| AMI Personal Semantic Schedule (childhood) | 2 | 52 | 0.47 | 0.41 (−0.10, 1.04) | 0.34 | 0.59 | 3 | 67 | 60 116 |
| Selective Reminding Test (free recall) | 2 | 46 | −0.47 | 0.44 (−1.09, 0.14) | −0.48 | −0.47 | 3 | 67 | 60 121 |
| AMI Autobiographical Incidents Schedule (childhood) | 2 | 52 | 0.37 | 0.41 (−0.20, 0.94) | 0.17 | 0.59 | 2 | 73 | 60 116 |
| AMI Personal Semantic Schedule (recent life) | 2 | 52 | −0.35 | 0.43 (−0.94, 0.24) | −1.25 | 0.45 | 1 | 73 | 60 116 |
| AMI Personal Semantic Schedule (early adulthood) | 2 | 52 | 0.34 | 0.41 (−0.23, 0.90) | 0.33 | 0.35 | 1 | 79 | 60 116 |
| CVLT (trial 4) | 3 | 159 | 0.30 | 0.30 (−0.03, 0.63) | −0.10 | 0.36 | 2 | 79 | 93 97 98 |
| AMI Autobiographical Incidents Schedule (recent life) | 2 | 52 | −0.25 | 0.42 (−0.83, 0.33) | −0.91 | 0.39 | 0 | 79 | 60 116 |
| CVLT (recognition discrimination) | 2 | 67 | −0.22 | 0.36 (−0.71, 0.27) | −0.46 | 0.54 | 0 | 85 | 87 93 |
| WMS Logical Memory (score not specified) | 4 | 125 | −0.20 | 0.41 (−0.59, 0.20) | −0.59 | 0.41 | 0 | 85 | 29 60 89 130 |
| Shopping List (delayed recall) | 2 | 68 | 0.18 | 0.37 (−0.34, 0.69) | 0.04 | 0.41 | 0 | 85 | 31 86 |
| Recognition Memory Test (faces) | 4 | 113 | 0.18 | 0.39 (−0.21, 0.56) | −0.33 | 1.06 | 0 | 85 | 28 82 119 120 |
| AMI Autobiographical Incidents Schedule (early adulthood) | 2 | 52 | −0.16 | 0.41 (−0.73, 0.40) | −0.28 | −0.04 | 0 | 85 | 60 116 |
| Benton Visual Retention Test | 3 | 194 | −0.16 | 0.29 (−0.49, 0.17) | −0.31 | 0.45 | 0 | 85 | 44 60 74 |

ACE, Addenbrooke's Cognitive Examination; AMI, Autobiographical Memory Interview; AVLT, Auditory Verbal Learning Test; CERAD−NP, Consortium to Establish a Registry for Alzheimer's Disease–Neuropsychological Battery; CVLT, California Verbal Learning Test; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; 95% CI, 95% CI for means; WMS, Wechsler Memory Scale (all editions).

*Effect sizes weighted by the inverse variance.

Overall, the effect sizes ranged from a minimum of −0.05 for the Judgement of Line Orientation Test to a maximum of 1.39 for the Graded Naming Test. Thus there was considerable variation in the extent to which the AD and FTD groups differed on the measures used by the 94 studies that were included in the current meta-analysis. This is also reflected in the per cent overlap statistics, which indicate that there was 100% overlap between the AD and FTD groups on the least discriminating measure (Judgement of Line Orientation Test) and 32% overlap between the two groups for the most discriminating measure (Graded Naming Test). Also interesting is the fact that the cognitive tests reported here were used by between two (eg, Addenbrooke's Cognitive Examination (ACE) subtests, Western Aphasia Battery (WAB) Comprehension) and 43 studies (Category Fluency). Thus while some measures have been used infrequently, others have been studied very extensively (eg, Category Fluency, Letter Fluency, Rey Complex Figure, Boston Naming Test). In terms of the statistical significance of the effect sizes, as indicated by CIs that do not span zero, there were 69 effect sizes that differed significantly from zero.

Of the 18 measures of orientation and attention, only the ACE orientation score met the study criteria (see table 3). This measure had a large and significant effect size, reflecting 41% overlap in the scores of the AD and FTD groups, with the AD group performing more poorly than the FTD group on this measure. The Nfs for this measure was also well in excess of the number of published studies using this measure, suggesting a high degree of confidence in this finding.

Only five measures of perception were used by two or more studies and these were all taken from the Visual Object and Space Perception battery (see table 4). None of these measures produced large effects, although position discrimination and dot counting revealed moderate group differences and CIs that did not span zero. However, these measures were associated with more than 50% overlap between the two groups and both had modest Nfs, indicating that the effect sizes could be reduced to a small effect size if there were 5–7 unpublished studies that had found small group differences. Therefore, none of the measures of perception accurately discriminated between persons with AD and FTD.

Memory was one of the most commonly assessed cognitive domains. In particular, the Rey Complex Figure Test and the Logical Memory subtest from the Wechsler Memory Scale (WMS) were very commonly used (see table 5). The recognition and delayed recall scores of the Auditory Verbal Learning Test (AVLT), the delayed recall trial of the Rey Complex Figure task, the Recognition Memory Test (words), the retention and delayed recall scores of the Logical Memory subtest from the WMS, the memory score from the ACE, the delayed recall score of the Visual Reproduction subtest from the WMS and the total recall score from the Selective Reminding Test all had large and significant effect sizes and less than 50% overlap between groups. Persons with AD performed more poorly than those with FTD on all of these measures. The delayed recall of the Rey Complex Figure task, in particular, has been used widely (21 studies) and the Nfs suggest that it is unlikely that this result would be negated by non-significant unpublished findings. The recognition score of the AVLT, the Recognition Memory Test for words, the retention score of the Logical Memory subtest from the WMS and the memory score of the ACE have been less widely used and, while acceptable, their Nfs were substantially lower.

Regarding tests of verbal abilities and language, there were more tests that fell into this category than any other. Eight of these test scores yielded large effect sizes, 95% CIs that did not

**Table 6** Verbal abilities and language: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N studies | N participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min d | Max d | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| Graded Naming Test | 3 | 88 | 1.39 | 0.44 (0.89, 1.88) | 0.61 | 3.01 | 18 | 32 | 28 78 109 |
| Word–Picture Matching | 5 | 123 | 1.05 | 0.44 (0.66, 1.44) | 0.37 | 2.00 | 21 | 41 | 30 79 109 119 120 |
| WAB Spontaneous Speech (fluency) | 2 | 227 | 1.04 | 0.20 (0.76, 1.32) | 0.48 | 1.29 | 8 | 45 | 94 124 |
| Pyramids and Palm Trees (words) | 5 | 185 | 1.02 | 0.36 (0.71, 1.34) | 0.65 | 1.72 | 21 | 45 | 82 90 114 119 120 |
| Other picture naming tasks | 11 | 291 | 0.99 | 0.45 (0.73, 1.26) | −0.45 | 2.93 | 43 | 45 | 30, 60, 82, 88, 109, 110, 112, 115, 119, 133, 136 |
| Pyramids and Palm Trees (pictures) | 4 | 116 | 0.90 | 0.40 (0.51, 1.30) | 0.29 | 1.48 | 14 | 48 | 78 82 119 120 |
| WAB Comprehension | 2 | 35 | 0.88 | 0.51 (0.18, 1.59) | 0.42 | 1.27 | 7 | 48 | 75 111 |
| WAB Repetition | 3 | 247 | 0.81 | 0.23 (0.55, 1.07) | 0.32 | 1.28 | 9 | 53 | 75 94 124 |
| Pyramids and Palm Trees (score not specified or average of words and pictures) | 5 | 123 | 0.79 | 0.43 (0.41, 1.17) | −0.39 | 1.31 | 15 | 53 | 28 30 80 115 136 |
| WAB Naming (object naming) | 3 | 247 | 0.78 | 0.23 (0.51, 1.04) | 0.30 | 1.11 | 9 | 53 | 75 94 124 |
| WAB Aphasia Quotient | 2 | 81 | 0.77 | 0.33 (0.31, 1.22) | 0.55 | 1.61 | 6 | 53 | 75 94 |
| WAB Praxis | 2 | 158 | 0.71 | 0.23 (0.39, 1.04) | 0.69 | 0.88 | 5 | 57 | 75 124 |
| WAB Naming (responsive speech) | 3 | 247 | 0.70 | 0.23 (0.44, 0.96) | 0.33 | 0.85 | 7 | 57 | 75 94 124 |
| WAB Naming (sentence completion) | 3 | 247 | 0.65 | 0.23 (0.39, 0.91) | 0.30 | 0.79 | 7 | 57 | 75 94 124 |
| WAIS Vocabulary | 3 | 325 | 0.62 | 0.30 (0.28, 0.96) | 0.40 | 0.77 | 6 | 62 | 62 86 94 |
| WAB Reading and writing (writing) | 2 | 152 | 0.62 | 0.24 (0.29, 0.95) | 0.47 | 0.64 | 4 | 62 | 75 124 |
| WAB Comprehension (yes/no) | 2 | 227 | 0.57 | 0.14 (0.30, 0.84) | 0.40 | 0.63 | 4 | 62 | 94 124 |
| WAB Spontaneous Speech (content) | 2 | 227 | 0.57 | 0.19 (0.30, 0.84) | 0.54 | 0.64 | 4 | 62 | 94 124 |
| ACE Naming | 2 | 170 | 0.55 | 0.24 (0.21, 0.88) | 0.36 | 0.76 | 3 | 62 | 85 102 |
| Letter Fluency | 39 | 1494 | 0.53 | 0.36 (0.42, 0.65) | −0.71 | 2.19 | 65 | 67 | 25, 27, 29−31, 42, 56, 60, 64, 65, 67−69, 72, 74, 76, 78, 80, 84, 86, 88, 92, 93, 96−98, 100, 101, 110, 111, 114, 118−121, 123, 126, 129 |
| WAB Reading and Writing (reading) | 2 | 171 | 0.49 | 0.22 (0.18, 0.80) | 0.30 | 0.51 | 3 | 67 | 75 124 |
| WAB Comprehension (word recognition) | 2 | 227 | 0.47 | 0.19 (0.21, 0.74) | 0.34 | 0.53 | 3 | 67 | 94 124 |
| ACE Language | 2 | 170 | 0.39 | 0.24 (0.06, 0.72) | 0.12 | 0.70 | 2 | 73 | 85 102 |
| Verb Similarity Task | 2 | 69 | 0.39 | 0.35 (−0.10, 0.87) | 0.30 | 0.51 | 2 | 73 | 108 114 |
| WAIS Information | 3 | 330 | 0.35 | 0.28 (0.04, 0.67) | 0.31 | 0.52 | 2 | 73 | 60 62 94 |
| Sentence Comprehension | 4 | 221 | 0.32 | 0.28 (0.04, 0.60) | −0.44 | 0.54 | 2 | 79 | 93 94 97 98 |
| ACE Verbal fluency (letter and category fluency) | 2 | 170 | 0.29 | 0.24 (−0.04, 0.62) | 0.19 | 0.42 | 1 | 79 | 85 102 |
| WAB Construction (calculation) | 2 | 162 | 0.26 | 0.23 (−0.05, 0.58) | −0.26 | 0.34 | 1 | 79 | 75 124 |
| Oral Praxis | 2 | 74 | 0.19 | 0.33 (−0.27, 0.65) | −0.17 | 0.83 | 0 | 85 | 68 123 |
| Boston Naming Test | 22 | 1352 | 0.17 | 0.30 (0.04, 0.29) | −2.24 | 1.11 | 0 | 85 | 25, 27, 31, 40, 44, 56, 68, 70, 74, 80, 86, 93, 97, 98, 105, 111, 114, 118, 126−128, 134 |
| Token Test | 4 | 182 | 0.17 | 0.32 (−0.15, 0.49) | −0.47 | 0.78 | 0 | 85 | 29 68 74 129 |
| WAB Construction (drawing) | 2 | 151 | −0.16 | 0.23 (−0.49, 0.16) | −0.17 | −0.10 | 0 | 85 | 75 124 |
| Category Fluency | 43 | 2481 | 0.15 | 0.30 (0.06, 0.24) | −5.44 | 1.35 | 0 | 85 | 25, 28, 30, 31, 40, 44, 60, 68−72, 75, 76, 80, 82, 86, 88, 92−94, 96−98, 100, 101, 105, 110−112, 115, 118−120, 123, 124, 126−131, 134 |
| Phrase Repetition | 3 | 158 | −0.07 | 0.30 (−0.40, 0.26) | −0.66 | 0.11 | 0 | 92 | 93 97 98 |
| Calculations | 4 | 227 | −0.06 | 0.28 (−0.33, 0.22) | −0.82 | 0.14 | 0 | 92 | 25 93 97 98 |

ACE, Addenbrooke's Cognitive Examination; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; WAB, Western Aphasia Battery; WAIS, Wechsler Adult Intelligence Scale (all editions); 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

span zero and large Nfs statistics, indicating that there were large and significant group differences on these measures and that these findings are unlikely to be negated by unpublished findings (see table 6). Specifically, these measures were: the Graded Naming Test, Word–Picture Matching, WAB Spontaneous Speech (fluency), the Pyramids and Palm Trees (word score, picture score), other picture naming tasks, WAB Comprehension and WAB Repetition. For all eight verbal measures, the FTD groups performed more poorly than the AD group. In addition to these eight measures, there were a number of other commonly used measures, such as the Pyramids and Palm Trees (score not specified) and letter fluency tasks, which showed moderate effect sizes, had 95% CIs that did not span zero and large Nfs statistics. However, because the effect sizes for these measures were relatively low (indicating small to medium differences in the AD and FTD means), there was an unacceptably high degree of overlap (53% and 67%, respectively) between the scores for the AD and FTD groups.

Of the seven measures of construction that were used by two or more studies, only the Beery Developmental Test of Visual Motor Integration demonstrated a large group difference, less than 50% overlap in scores, 95% CIs that did not span zero and a satisfactory Nfs (see table 7). As with the abovementioned memory tests, the effect size was negative, indicating that the AD groups performed more poorly on this task than the FTD groups.

Despite researchers using a number of different measures of concept formation and reasoning, none of them showed large group differences (ie, *d* >0.8) (see table 8). Indeed, there was between 57% and 92% overlap in the scores of the AD and FTD groups on all 12 measures. Commonly used measures, such as the Wisconsin Card Sorting Test and Raven's Progressive Matrices, did not successfully discriminate between AD and FTD when the findings of all studies that had used these measures were considered, although the minimum and maximum effects sizes for these tests indicate that there were individual studies that reported large effects for these tests.

When the classifications used by Lezak *et al*[45] and Spreen and Strauss[46] were used to group the cognitive tests, there were only four tests that fell into the category of executive function and motor performance (see table 9), although other tests that are

**Table 7** Construction: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N studies | N participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min *d* | Max *d* | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| Beery Developmental Test of Visual Motor Integration | 2 | 68 | −0.95 | 0.39 (−1.49, −0.41) | −1.00 | −0.86 | 8 | 45 | 31 86 |
| ACE Visuospatial | 2 | 170 | −0.58 | 0.24 (−0.91, −0.25) | −0.65 | −0.50 | 4 | 62 | 85 102 |
| Clock drawing tests | 4 | 392 | −0.53 | 0.24 (−0.76, −0.29) | −1.41 | −0.06 | 7 | 67 | 69 71 92 113 |
| Rey Complex Figure (copy) | 25 | 874 | −0.43 | 0.36 (−0.57, −0.29) | −1.47 | 0.52 | 29 | 73 | 25, 27–31, 42, 68, 78, 86, 88–90, 93, 97, 98, 110, 115, 119, 120, 123, 129, 130, 134 |
| Other figure copying tasks | 7 | 551 | −0.37 | 0.24 (−0.55, −0.19) | −0.77 | 0.41 | 6 | 73 | 40, 44, 69, 70, 92, 105, 126 |
| WAIS Block Design | 6 | 443 | −0.29 | 0.32 (−0.54, −0.03) | −0.84 | 0.29 | 3 | 79 | 56, 62, 74, 86, 94, 124 |
| WAIS Object Assembly | 2 | 300 | 0.28 | 0.26 (−0.08, 0.63) | 0.21 | 0.42 | 1 | 79 | 62 94 |

ACE, Addenbrooke's Cognitive Examination; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; WAIS, Wechsler Adult Intelligence Scale (all editions); 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

frequently considered to be executive tasks have been included in other categories (eg, verbal fluency tasks, Wisconsin Card Sorting Test, Wechsler Adult Intelligence Scale-Similarities).[137 138] None of these met the criteria specified for discriminating between AD and FTD.

Finally, of the five measures of general ability and intelligence that were used by the studies included in this meta-analysis, the MMSE was the only measure that revealed a large negative effect (see table 10). This indicates that AD participants were more impaired on this measure than FTD participants. Despite its widespread use, only those studies (n = 5) that had pathological confirmation of a patient's diagnosis were used in the calculation of this effect size. This was done in order to avoid a situation where the MMSE was being used both in group allocation (independent variable) and as a measure of cognitive performance (dependent variable).

## Moderator variables

The influence of four methodological variables (age, years since diagnosis, education, diagnostic criteria) on the effect sizes was also examined using Pearson *r* correlation coefficients and a WLS multiple regression analysis. ''Diagnostic criteria'' refer to the rating given to the quality of the diagnostic criteria that was used to diagnose AD and FTD (ie, pathological confirmation, published criteria or other criteria). The mean age, years since diagnosis, education and diagnostic criteria scores of those studies that reported this data were firstly correlated with the weighted mean effect sizes for these studies. Small and non-significant correlations were observed for age (r = −0.01, n = 90, p = 0.94), years since diagnosis (r = −0.10, n = 38, p = .57), education (r = 0.11, n = 63, p = 0.41) and diagnostic criteria (r = −0.08, n = 94, p = 0.46), indicating that these variables were not significantly related to the effect sizes. A WLS multiple regression was additionally performed on the data from the 22 studies that reported data for all four variables. The weighted mean effect size for a study was the dependent variable, the four moderator variables were the independent variables, and the inverse variance was used as the weighting variable. The results of this regression analysis are presented in table 11. The final model was non-significant (p = 0.98) and accounted for only 23% of the variance. Therefore, both the correlations and the WLS multiple regression indicate that age, years since diagnosis, education and diagnostic criteria did not significantly contribute to the effect sizes reported in this study. Thus there did not appear to be any systematic variation in the performance of the AD and FTD groups on the different tests of cognitive ability as a consequence of between study differences in these variables, suggesting that it is acceptable to combine the results of studies that differed on these methodological variables.

## DISCUSSION

Overall, the data for this meta-analysis were obtained from 94 studies that examined the cognitive performance of 2936

**Table 8** Concept formation and reasoning: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N Studies | N Participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min *d* | Max *d* | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| Proverb Interpretation | 2 | 111 | 0.66 | 0.29 (0.26, 1.05) | 0.25 | 1.00 | 5 | 57 | 69 92 |
| WAIS Comprehension | 4 | 359 | 0.65 | 0.31 (0.35, 0.95) | 0.40 | 0.80 | 9 | 57 | 62 72 86 94 |
| Cognitive Estimation Test | 2 | 94 | −0.51 | 0.30 (−0.93, −0.09) | −0.67 | −0.16 | 3 | 67 | 63 100 |
| Raven's Progressive Matrices | 10 | 686 | −0.47 | 0.28 (−0.64, −0.30) | −1.75 | 0.29 | 14 | 67 | 29, 44, 68, 74, 96, 101, 123, 124, 129, 130 |
| WAIS Similarities | 6 | 435 | 0.37 | 0.31 (0.12, 0.62) | −0.08 | 0.78 | 5 | 73 | 60 62 72 86 94 126 |
| WCST (errors) | 3 | 109 | −0.30 | 0.34 (−0.68, 0.08) | −0.56 | −0.20 | 1 | 79 | 31 110 123 |
| WAIS Arithmetic | 2 | 306 | 0.29 | 0.25 (−0.06, 0.63) | −0.22 | 0.59 | 1 | 79 | 62 94 |
| WAIS Picture Completion | 2 | 308 | 0.25 | 0.25 (−0.10, 0.59) | 0.08 | 0.53 | 0 | 79 | 62 94 |
| WCST (perseverations) | 12 | 326 | 0.19 | 0.41 (−0.05, 0.42) | −1.16 | 1.39 | 0 | 85 | 29, 31, 68, 74, 76, 78, 86, 88, 100, 110, 121, 131 |
| WCST (categories) | 13 | 367 | 0.11 | 0.40 (−0.11, 0.33) | −1.41 | 1.83 | 0 | 92 | 28, 29, 31, 68, 74, 76, 78, 88, 100, 110, 121, 123, 131 |
| WAIS Picture Arrangement | 4 | 479 | −0.07 | 0.26 (−0.32, 0.19) | −0.44 | 0.25 | 0 | 92 | 44 62 86 94 |
| Judgement of Line Orientation | 2 | 42 | −0.05 | 0.46 (−0.68, 0.59) | −0.26 | 0.30 | 0 | 92 | 30 111 |

Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; WAIS, Wechsler Adult Intelligence Scale (all editions); WCST, Wisconsin Card Sorting Test; 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

**Table 9** Executive and motor functions: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N Studies | N Participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min *d* | Max *d* | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| FAB | 4 | 314 | 0.43 | 0.25 (0.19, 0.67) | −0.12 | 1.29 | 5 | 73 | 59 104 126 129 |
| Frontal Composite Error Score | 2 | 143 | 0.29 | 0.26 (−0.07, 0.64) | 0.25 | 0.33 | 1 | 79 | 97 98 |
| FAB (go no go) | 2 | 170 | −0.13 | 0.24 (−0.46, 0.20) | −0.50 | 0.56 | 0 | 92 | 126 129 |
| Design Fluency | 7 | 306 | −0.06 | 0.32 (−0.30, 0.17) | −0.51 | 0.61 | 0 | 92 | 25 31 86 93 97 98 131 |

FAB, Frontal Assessment Battery; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

persons with AD and 1748 persons with FTD. Forty-seven cognitive tests, which yielded 115 different scores, were used by more than one study. While the participants with AD were, on average, older and less educated than the FTD participants, the two groups were comparable in terms of time since diagnosis. The lower age of the FTD group is consistent with the fact that FTD tends to have an earlier onset.[14] FTD participants also had higher MMSE scores than the AD participants. However, this difference in MMSE scores is likely to reflect the fact that this measure was originally designed to detect the deficits associated with AD (eg, memory, praxis) rather than FTD.

The current meta-analysis found that a wide variety of tests have been used in research that has examined the cognitive deficits associated with AD and FTD and that there is considerable variation in the ability of these measures to distinguish between AD and FTD. However, in order for a cognitive test to be useful for differential diagnosis, it was argued that it must be able to distinguish between the performance of these two groups (indicated by large effects and small %OL). There should also be a high degree of confidence in the accuracy with which the measure distinguishes between the two groups (measured by the 95% CI) and we need to be assured that the conclusions drawn from the research literature are not systematically biased by the tendency to publish statistically significant findings (Nfs statistic). When these three criteria were applied to all of the measures analysed in this study, there were only 19 out of 115 measures that met these criteria. More specifically, persons with AD performed more poorly than those with FTD on 12 of the 19 measures, all of which assessed orientation, memory or general cognitive ability. These tests were (in order of discriminative ability): the ACE orientation subtest, the MMSE, the Beery Developmental Test of Visual Motor Integration and nine measures of memory (AVLT recognition and delayed recall scores, Rey Complex Figure delayed recall score, Recognition Memory Test (words), WMS Logical Memory per cent retention and delayed recall scores, ACE memory subtest, WMS Visual Reproduction delayed recall score and Selective Reminding Test total recall).

In contrast, persons with FTD performed more poorly than those with AD on seven measures of verbal ability (Graded Naming Test, Word–Picture Matching, WAB Spontaneous Speech (fluency), Pyramids and Palm Trees (word score, picture score), other picture naming tasks and WAB Comprehension). Importantly, however, there was still between 32% and 48% overlap in the scores of the AD and FTD groups on all 19 of these measures. This level of overlap suggests that, even when the most discriminating cognitive tests are used, the differential diagnosis of AD and FTD remains problematic. Failure to find tests that clearly discriminate between AD and FTD confirms previous research which found that the performance of AD and FTD participants did not differ significantly on a large range of tests.[27 44 68 72 94 130]

The finding that there were nine measures of memory that were among the most useful for differentiating between AD and FTD is consistent with research and clinical observations that memory is differentially affected in AD and FTD. Numerous previous studies have reported that AD patients experience greater memory problems than those with FTD.[25 27 28 30 44 66 68 72 78 84 85 90 91 94 112] In contrast, persons diagnosed with FTD had more difficulty with tests of verbal ability and language than those with AD, as demonstrated by eight large effect sizes in this domain. This supports previous research which indicated that FTD is associating with language impairment.[5 7 9] In addition, numerous previous studies have reported that persons with FTD have more difficulty with executive tasks than AD patients.[2–5 9 11 12 27 67] While executive tasks would be expected to discriminate between AD and FTD, given the frontal pathology associated with FTD,[9–11] this meta-analysis did not identify any measures of executive functioning that adequately distinguished between the two groups. Moreover, other tests that are thought to draw on executive functions for successful completion,[137–139] such as verbal fluency, Trails and the Stroop tests, failed to meet the current study criteria.

While there was considerable variation in the effect sizes for different studies, between study variations in the age of

**Table 10** General ability and intelligence: weighted Cohen's *d* effect sizes* for each test, rank ordered by size

| | N Studies | N Participants | Mean Cohen's $d_w$ | SD $d_w$ (95% CI) | Min *d* | Max *d* | Nfs | %OL | Study references |
|---|---|---|---|---|---|---|---|---|---|
| MMSE | 5 | 168 | −0.98 | 0.37 (−1.31, −0.65) | −1.29 | −0.01 | 19 | 45 | 54–58 |
| WAIS Verbal IQ | 5 | 427 | 0.51 | 0.29 (0.26, 0.77) | 0.14 | 0.92 | 8 | 67 | 31 62 83 94 132 |
| WAIS Full Scale IQ | 6 | 433 | 0.30 | 0.31 (0.05, 0.56) | −0.23 | 0.62 | 3 | 79 | 31 62 74 83 86 94 |
| ACE | 4 | 171 | −0.18 | 0.34 (−0.51, 0.14) | −0.70 | 0.57 | 0 | 85 | 78 99 102 115 |
| WAIS Performance IQ | 5 | 428 | 0.16 | 0.29 (−0.09, 0.41) | −0.56 | 0.69 | 0 | 85 | 31 62 83 94 132 |

ACE, Addenbrooke's Cognitive Examination; Max *d*, maximum effect size; Mean $d_w$, mean weighted effect size; Min *d*, minimum effect size; MMSE, Mini-Mental State Examination; N, number of studies contributing to the effect size; Nfs, fail safe N; %OL, per cent overlap between the Alzheimer's disease and frontotemporal dementia groups; SD $d_w$, SD of weighted effect size; WAIS, Wechsler Adult Intelligence Scale (all editions); 95% CI, 95% CI for means.
*Effect sizes weighted by the inverse variance.

**Table 11** Weighted least squares multiple regression results for moderator variables predicting effect size

| Variable | β | SE (95% CI) |
|---|---|---|
| Age | 0.00 | 0.06 (−0.114, 0.122) |
| Years since diagnosis | −0.05 | 0.20 (−0.454, 0.350) |
| Years of education | 0.03 | 0.08 (−0.116, 0.184) |
| Diagnostic criteria | −0.11 | 0.39 (−0.880, 0.660) |

β, estimate of the unstandardised regression coefficient; SE, standard error of regression coefficient.

participants, years since diagnosis, years of education and the diagnostic criteria that were used were not related to the size of the effects reported by these studies. The finding that years since diagnosis was not related to the group differences is surprising given that the symptoms associated with both AD and FTD vary with the stage of illness, with cognitive performance declining over time.[4][11][12][43][140–142] However, demographic variables such as age, gender, years since diagnosis and education were not consistently reported. For example, only approximately 40% of the studies reported years since diagnosis. Thus there is the potential for significant but unreported variation in these variables. Also important is the fact that mean effect sizes were calculated for each study in order to complete these analyses. Thus the effect sizes for all of the individual measures that were used by a given study were averaged. If a study yielded both small and large effect sizes for different tests, these were combined to calculate a mean effect size, which was then correlated with the moderator variables. A more useful analysis would involve analysing each measure separately by correlating the effect sizes from all of the studies that used a particular measure with the age, education and years since diagnosis data for those studies. However, this was not possible as the maximum number of studies that used any given measure was 43. When those studies that did not provide data for the moderator variables were excluded, the sample size was reduced to an unacceptably low number (ie, 22). Thus the current moderator analysis was necessarily limited. Similarly, current medications and whether depressed participants were excluded was rarely reported. Therefore, these variables could not be examined, despite the potential for medications and depression to affect cognitive performance. It is possible that these factors contributed to the effect sizes. It is, therefore, essential that this information is included in future publications of primary research to allow an accurate and detailed synthesis of the research findings.

There are a number of limitations to the current meta-analysis that warrant consideration. Firstly, rating inventories were not included in the current meta-analysis as they are not classified as objective cognitive tests.[143][144] However, given the behavioural nature of FTD, a meta-analysis of these behavioural rating scales may yield useful findings. In addition, assessments of social cognition, personality and empathy measures, which have been recommended in the differential diagnosis of dementia,[11][13][64][145] did not fall within the scope of the current meta-analysis but may also be worthy of consideration.

Secondly, it is important to note that the current meta-analysis specifically examined AD and FTD. The overlapping cognitive features of these types of dementia are highlighted by the fact that few tests clearly discriminated between these two groups. While greater differences would be expected if the cognitive performance of the AD and FTD groups were compared with healthy controls, such an analysis would not adequately address the question of differential diagnosis.

Thirdly, FTD is described using a variety of terms, which were combined in the current meta-analysis. This may have increased the variability in the test performance of the FTD group. Although it would be desirable to analyse these subgroups separately, the labels and the diagnostic criteria that are applied to these subgroups are not used consistently, thereby precluding a more specific analysis of potential subtypes. Moreover, more research is needed on each of these subgroups if they are to be analysed separately.

While this analysis indicates that many test scores do not clearly discriminate between AD and FTD, there is some evidence that qualitative aspects of performance may better distinguish the two groups. For example, a recent study indicated that the consideration of performance characteristics and specific error types increased the accuracy of the differential diagnosis of AD and FTD.[33] However, given that these aspects of performance are often not quantified, discrimination between AD and FTD is largely reliant on subjective clinical observations.

Finally, the effect sizes derived from measures used by single studies were not included in the current meta-analysis. Technically, a meta-analysis requires two or more studies in order to aggregate primary research findings.[41] One consequence of this is that there may be interesting and innovative measures that discriminate between groups but are not in widespread use and were not included in the current analysis.

## CONCLUSIONS

In summary, the findings of this meta-analysis suggest that the neuropsychological tests that best discriminate between AD and FTD are ACE orientation, ACE memory, recognition and delayed recall of the AVLT, delayed recall of the Rey Complex Figure, the words version of the Recognition Memory Test, WMS Logical Memory (per cent retention and delayed recall), WMS Visual Reproduction Test (delayed), total recall of the Selective Reminding Test, Graded Naming Test, Word–Picture Matching, WAB Spontaneous Speech (fluency), Pyramids and Palm Trees (word score, picture score), picture naming tasks, WAB Comprehension, the Beery Developmental Test of Visual Motor Integration and MMSE. However, it is important to note that none of the tests showed acceptably low overlap between the scores of the two groups to confidently make a differential diagnosis. Therefore, these cognitive tests must be used cautiously and in conjunction with other diagnostic information, such as medical history, behavioural observations, imaging and information from relatives when making a differential diagnosis.

. . . . . . . . . . . . . . . . . . . . . .

**Authors' affiliations**
**A D Hutchinson, J L Mathias,** School of Psychology, University of Adelaide, Adelaide, Australia

Competing interests: None.

## REFERENCES

1 **Prigerson HG**. Costs to society of family caregiving for patients with end-stage Alzheimer's disease. *N Engl J Med* 2003;**349**:1891–12.
2 **Davidoff DA**. The recognition and treatment of dementia. In: Ellison JM, Weinstein C, Hodel-Malinofsky T, eds. *The psychotherapist's guide to neuropsychiatry:Diagnostic and treatment issues*. Washington, DC: American Psychiatric Association, 1994:487–556.

3　**Moss MB**, Albert MS. Neuropsychology of Alzheimer's disease. In: White RF, eds. *Clinical syndromes in adult neuropsychology:the practitioner's handbook.* New York: Elsevier Science, 1992:305–43.

4　**Usman MA**. Frontotemporal dementias. In: Nussbaum PD, eds. *Handbook of neuropsychology and aging.* New York, NY: Plenum Press, 1997:159–76.

5　**Norris MP**, Haines ME. Psychological and neuropsychological aspects of Lewy body and frontal dementia. In: Mellow AM, eds. *Handbook of dementia:psychological, neurological, and psychiatric perspectives.* Hoboken, NJ, US: John Wiley & Sons, 2003:115–48.

6　**Duara R**, Barker W, Luis CA. Frontotemporal dementia and Alzheimer's disease: differential diagnosis. *Dement Geriatr Cogn Disord* 1999;**10**(suppl 1):37–42.

7　**Pasquier F**. Early diagnosis of dementia: neuropsychology. *J Neurol* 1999;**246**:6–15.

8　**Greicius MD**, Geschwind MD, Miller BL. Presenile dementia syndromes: an update on taxonomy and diagnosis. *J Neurol Neurosurg Psychiatry* 2002;**72**:691–700.

9　**Hou CE**, Carlin D, Miller BL. Non-Alzheimer's disease dementias: anatomic, clinical, and molecular correlates. *Can J Psychiatry* 2004;**49**:164–71.

10　**Yeaworth RC**, Burke WJ. Frontotemporal dementia: a different kind of dementia. *Arch Psychiatr Nurs* 2000;**14**:249–53.

11　**Miller BL**, Diehl J, Freedman M, *et al.* International approaches to frontotemporal dementia diagnosis: from social cognition to neuropsychology. *Ann Neurol* 2003;**54**(Suppl 5):S7–10.

12　**Moss MB**, Albert MS, Kemper TL. Neuropsychology of frontal lobe dementia. In: White RF, eds. *Clinical syndromes in adult neuropsychology:the practitioner's handbook.* New York, NY: Elsevier Science, 1992:287–303.

13　**Rankin KP**, Baldwin E, Pace-Savitsky C, *et al.* Self awareness and personality change in dementia. *J Neurol Neurosurg Psychiatry* 2005;**76**:632–9.

14　**Walker AJ**, Meares S, Sachdev PS, *et al.* The differentiation of mild frontotemporal dementia from Alzheimer's disease and healthy aging by neuropsychological tests. *Int Psychogeriatr* 2005;**17**:57–68.

15　**Pettenati C**, Annicchiarico R, Caltagirone C. Clinical pharmacology of anti-Alzheimer drugs. *Fundam Clin Pharmacol* 2003;**17**:659–72.

16　**Standridge JB**. Pharmacotherapeutic approaches to the treatment of Alzheimer's disease. *Clin Ther* 2004;**26**:615–30.

17　**Cummings JL**. Alzheimer's disease. *N Engl J Med* 2004;**351**:56–67.

18　**Mintzer JE**, Kershaw P. The efficacy of galantamine in the treatment of Alzheimer's disease: comparison of patients previously treated with acetylcholinesterase inhibitors to patients with no prior exposure. *Int J Geriatr Psychiatry* 2003;**18**:292–7.

19　**Raskind MA**, Peskind ER, Wessel T, *et al.* Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology* 2000;**54**:2261–8.

20　**Winblad B**, Engedal K, Soininen H, *et al.* A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology* 2001;**57**:489–95.

21　**Mohs RC**, Doody RS, Morris JC, *et al.* A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology* 2001;**57**:481–8.

22　**Cummings JL**. Use of cholinesterase inhibitors in clinical practice: evidence-based recommendations. *Am J Geriatr Psychiatry* 2003;**11**:131–45.

23　**Adler G**, Teufel M, Drach LM. Pharmacological treatment of frontotemporal dementia: treatment response to the MAO-A inhibitor moclobemide. *Int J Geriatr Psychiatry* 2003;**18**:653–5.

24　**Rascovsky K**, Salmon DP, Ho GJ, *et al.* Cognitive profiles differ in autopsy-confirmed frontotemporal dementia and AD. *Neurology* 2002;**58**:1801–8.

25　**Kramer JH**, Jurik J, Sha SJ, *et al.* Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cogn Behav Neurol* 2003;**16**:211–18.

26　**Minami M**, Kimura M, Iwamoto N, *et al.* Endothelin-1-like immunoreactivity in cerebral cortex of Alzheimer-type dementia. *Prog Neuropsychopharmacol Biol Psychiatry* 1995;**19**:509–13.

27　**Pachana NA**, Boone KB, Miller BL, *et al.* Comparison of neuropsychological functioning in Alzheimer's disease and frontotemporal dementia. *J Int Neuropsychol Soc* 1996;**2**:505–10.

28　**Perry RJ**, Hodges JR. Differentiating frontal and temporal variant frontotemporal dementia from Alzheimer's disease. *Neurology* 2000;**54**:2277–84.

29　**Frisoni GB**, Pizzolato G, Geroldi C, *et al.* Dementia of the frontal type: neuropsychological and [99Tc]-HM-PAO SPET features. *J Geriatr Psychiatry Neurol* 1995;**8**:42–8.

30　**Hodges JR**, Patterson K, Ward R, *et al.* The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: a comparative neuropsychological study. *Neuropsychology* 1999;**13**:31–40.

31　**Razani J**, Boone KB, Miller BL, *et al.* Neuropsychological performance of right- and left-frontotemporal dementia compared to Alzheimer's disease. *J Int Neuropsychol Soc* 2001;**7**:468–80.

32　**Perri R**, Koch G, Carlesimo GA, *et al.* Alzheimer's disease and frontal variant of frontotemporal dementia—a very brief battery for cognitive and behavioural distinction. *J Neurol* 2005;**252**:1238–44.

33　**Thompson JC**, Stopford CL, Snowden JS, *et al.* Qualitative neuropsychological performance characteristics in frontotemporal dementia and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2005;**76**:920–7.

34　**Harciarek M**, Jodzio K. Neuropsychological differences between frontotemporal dementia and Alzheimer's disease: a review. *Neuropsychol Rev* 2005;**15**:131–45.

35　**Hedges LV**, Olkin I. *Statistical methods for meta-analysis.* New York: Academic, 1985.

36　**Rosenthal R**, DiMatteo MR. Meta-analysis: recent developments in quantitative methods for literature reviews. *Annu Rev Psychol* 2001;**52**:59–82.

37　**Cooper H**, Hedges LV, eds. *The handbook of research synthesis.* New York, NY: Russell Sage Foundation, 1994.

38　**Cohen J**. *Statistical power analyses for the behavioral sciences*, 2nd edn., New York: Academic Press 1988.

39　**Grossman M**, McMillan C, Moore P, *et al.* What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain* 2004;**127**:628–49.

40　**Gee J**, Ding L, Xie Z, *et al.* Alzheimer's disease and frontotemporal dementia exhibit distinct atrophy-behavior correlates: a computer-assisted imaging study. *Acad Radiol* 2003;**10**:1392–401.

41　**Rosenthal R**. Writing meta-analytic reviews. *Psychol Bull* 1995;**118**:183–92.

42　**Nestor PJ**, Graham KS, Bozeat S, *et al.* Memory consolidation and the hippocampus: further evidence from studies of autobiographical memory in semantic dementia and frontal variant frontotemporal dementia. *Neuropsychologia* 2002;**40**:633–54.

43　**McKhann G**, Drachman D, Folstein M, *et al.* 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* 1984;**34**:939–44.

44　**Binetti G**, Locascio JJ, Corkin S, *et al.* Differences between Pick disease and Alzheimer disease in clinical appearance and rate of cognitive decline. *Arch Neurol* 2000;**57**:225–32.

45　**Lezak MD**, Howieson DB, Loring DW, *et al.* Neuropsychological assessment, 4th edn., New York: Oxford University Press 2004.

46　**Spreen O**, Strauss E. *A compendium of neuropsychological tests: administration, norms, and commentary*, 2nd edn., New York, NY: Oxford University Press 1998.

47　**Zakzanis KK**. Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Arch Clin Neuropsychol* 2001;**16**:653–67.

48　**Lipsey MW**, Wilson DB. *Practical meta-analysis.* Thousand Oaks, CA, US: Sage Publications, 2001.

49　**Zakzanis KK**, Leach L, Kaplan E. *Neuropsychological differential diagnosis.* Lisse, Netherlands: Swets and Zeitlinger, 1999.

50　**The Lund and Manchester Groups**. Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 1994;**57**:416–18.

51　**Steel PD**, Kammeyer-Mueller JD. Comparing meta-analytic moderator estimation techniques under realistic conditions. *J Appl Psychol* 2002;**87**:96–111.

52　**Wall MM**, Li R. Comparison of multiple regression to two latent variable techniques for estimation and prediction. *Stat Med* 2003;**22**:3671–85.

53　**Hedges LV**. Fixed effects models. In: Cooper H, Hedges LV, eds. *The handbook of research synthesis.* New York: Russell Sage Foundation, 1994:285–99.

54　**Blass DM**, Hatanpaa KJ, Brandt J, *et al.* Dementia in hippocampal sclerosis resembles frontotemporal dementia more than Alzheimer disease. *Neurology* 2004;**63**:492–7.

55　**Varma AR**, Snowden JS, Lloyd JJ, *et al.* Evaluation of the NINCDS-ADRDA criteria in the differentiation of Alzheimer's disease and frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 1999;**66**:184–8.

56　**Knopman DS**, Christensen KJ, Schut LJ, *et al.* The spectrum of imaging and neuropsychological findings in Pick's disease. *Neurology* 1989;**39**:362–8.

57　**Lopez OL**, Litvan I, Catt KE, *et al.* Accuracy of four clinical diagnostic criteria for the diagnosis of neurodegenerative dementias. *Neurology* 1999;**53**:1292–9.

58　**Likeman M**, Anderson VM, Stevens JM, *et al.* Visual assessment of atrophy on magnetic resonance imaging in the diagnosis of pathologically confirmed young-onset dementias. *Arch Neurol* 2005;**62**:1410–16.

59　**Slachevsky A**, Villalpando JM, Sarazin M, *et al.* Frontal assessment battery and differential diagnosis of frontotemporal dementia and Alzheimer disease. *Arch Neurol* 2004;**61**:1104–7.

60　**Thomas-Anterion C**, Jacquin K, Laurent B. Differential mechanisms of impairment of remote memory in Alzheimer's and frontotemporal dementia. *Dement Geriatr Cogn Disord* 2000;**11**:100–6.

61　**Forstl H**, Besthorn C, Hentschel F, *et al.* Frontal lobe degeneration and Alzheimer's disease: a controlled study on clinical findings, volumetric brain changes and quantitative electroencephalography data. *Dementia* 1996;**7**:27–34.

62　**Yamashita H**, Hirono N, Ikeda M, *et al.* Examining the diagnostic utility of the Fuld cholinergic deficit profile on the Japanese WAIS-R. *J Clin Exp Neuropsychol* 1997;**19**:300–4.

63　**Mendez MF**, Doss RC, Cherrier MM. Use of the cognitive estimations test to discriminate frontotemporal dementia from Alzheimer's disease. *J Geriatr Psychiatry Neurol* 1998;**11**:2–6.

64　**Lindau M**, Almkvist O, Johansson SE, *et al.* Cognitive and behavioral differentiation of frontal lobe degeneration of the non-Alzheimer type and Alzheimer's disease. *Dement Geriatr Cogn Disord* 1998;**9**:205–13.

65　**Warkentin S**, Passant U. Functional imaging of the frontal lobes in organic dementia. Regional cerebral blood flow findings in normals, in patients with frontotemporal dementia and in patients with Alzheimer's disease, performing a word fluency test. *Dement Geriatr Cogn Disord* 1997;**8**:105–9.

66　**Souliez L**, Pasquier F, Lebert F, *et al.* Generation effect in short-term verbal and visuospatial memory: comparisons between dementia of Alzheimer type and dementia of frontal lobe type. *Cortex* 1996;**32**:347–56.

67　**Silveri MC**, Salvigni BL, Cappa A, *et al.* Impairment of verb processing in frontal variant-frontotemporal dementia: a dysexecutive symptom. *Dement Geriatr Cogn Disord* 2003;**16**:296–300.

68 **Siri S**, Benaglio I, Frigerio A, et al. A brief neuropsychological assessment for the differential diagnosis between frontotemporal dementia and Alzheimer's disease. Eur J Neurol 2001;**8**:125–32.

69 **Moretti R**, Torre P, Antonello RM, et al. Ten-Point Clock Test: a correlation analysis with other neuropsychological tests in dementia. Int J Geriatr Psychiatry 2002;**17**:347–53.

70 **Diehl J**, Kurz A. Frontotemporal dementia: patient characteristics, cognition, and behaviour. Int J Geriatr Psychiatry 2002;**17**:914–18.

71 **Meulen EF**, Schmand B, van Campen JP, et al. The seven minute screen: a neurocognitive screening test highly sensitive to various types of dementia. J Neurol Neurosurg Psychiatry 2004;**75**:700–5.

72 **Gregory CA**, Orrell M, Sahakian B, et al. Can frontotemporal dementia and Alzheimer's disease be differentiated using a brief battery of tests? Int J Geriatr Psychiatry 1997;**12**:375–83.

73 **Lindeboom J**, Schmand B, Tulner L, et al. Visual association test to detect early dementia of the Alzheimer type. J Neurol Neurosurg Psychiatry 2002;**73**:126–33.

74 **Starkstein SE**, Migliorelli R, Teson A, et al. Specificity of changes in cerebral blood flow in patients with frontal lobe dementia. J Neurol Neurosurg Psychiatry 1994;**57**:790–6.

75 **Karbe H**, Kertesz A, Polk M. Profiles of language impairment in primary progressive aphasia. Arch Neurol 1993;**50**:193–201.

76 **Cappa SF**, Binetti G, Pezzini A, et al. Object and action naming in Alzheimer's disease and frontotemporal dementia. Neurology 1998;**50**:351–5.

77 **Halpern C**, McMillan C, Moore P, et al. Calculation impairment in neurodegenerative diseases. J Neurol Sci 2003;**208**:31–8.

78 **Gregory C**, Lough S, Stone V, et al. Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. Brain 2002;**125**:752–64.

79 **Noble K**, Glosser G, Grossman M. Oral reading in dementia. Brain Lang 2000;**74**:48–69.

80 **Grossman M**, Smith EE, Koenig PL, et al. Categorization of object descriptions in Alzheimer's disease and frontotemporal dementia: limitation in rule-based processing. Cogn Affect Behav Neurosci 2003;**3**:120–32.

81 **Pasquier F**, Lebert F, Grymonprez L, et al. Verbal fluency in dementia of frontal lobe type and dementia of Alzheimer type. J Neurol Neurosurg Psychiatry 1995;**58**:81–4.

82 **Galton CJ**, Patterson K, Graham K, et al. Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. Neurology 2001;**57**:216–25.

83 **Kitagaki H**, Mori E, Hirono N, et al. Alteration of white matter MR signal intensity in frontotemporal dementia. AJNR Am J Neuroradiol 1997;**18**:367–78.

84 **Frisoni GB**, Laakso MP, Beltramello A, et al. Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. Neurology 1999;**52**:91–100.

85 **Mathuranath PS**, Nestor PJ, Berrios GE, et al. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. Neurology 2000;**55**:1613–20.

86 **Boone KB**, Miller BL, Swartz R, et al. Relationship between positive and negative symptoms and neuropsychological scores in frontotemporal dementia and Alzheimer's disease. J Int Neuropsychol Soc 2003;**9**:698–709.

87 **Glosser G**, Gallo JL, Clark CM, et al. Memory encoding and retrieval in frontotemporal dementia and Alzheimer's disease. Neuropsychology 2002;**16**:190–6.

88 **Piolino P**, Desgranges B, Belliard S, et al. Autobiographical memory and autonoetic consciousness: triple dissociation in neurodegenerative diseases. Brain 2003;**126**:2203–19.

89 **Frisoni GB**, Beltramello A, Geroldi C, et al. Brain atrophy in frontotemporal dementia. J Neurol Neurosurg Psychiatry 1996;**61**:157–65.

90 **Galton CJ**, Gomez-Anson B, Antoun N, et al. Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. J Neurol Neurosurg Psychiatry 2001;**70**:165–73.

91 **Pasquier F**, Grymonprez L, Lebert F, et al. Memory impairment differs in frontotemporal dementia and Alzheimer's disease. Neurocase 2001;**7**:161–71.

92 **Moretti R**, Torre P, Antonello RM, et al. Olanzapine as a treatment of neuropsychiatric disorders of Alzheimer's disease and other dementias: a 24-month follow-up of 68 patients. Am J Alzheimers Dis Other Demen 2003;**18**:205–14.

93 **Boxer AL**, Rankin KP, Miller BL, et al. Cinguloparietal atrophy distinguishes Alzheimer disease from semantic dementia. Arch Neurol 2003;**60**:949–56.

94 **Kertesz A**, Davidson W, McCabe P, et al. Behavioral quantitation is more sensitive than cognitive testing in frontotemporal dementia. Alzheimer Dis Assoc Disord 2003;**17**:223–9.

95 **Lavenu I**, Pasquier F. Perception of emotion on faces in frontotemporal dementia and Alzheimer's disease: a longitudinal study. Dement Geriatr Cogn Disord 2005;**19**:37–41.

96 **Maeshima S**, Osawa A, Maeshima E, et al. Usefulness of a cube-copying test in outpatients with dementia. Brain Inj 2004;**18**:889–98.

97 **Liu W**, Miller BL, Kramer JH, et al. Behavioral disorders in the frontal and temporal variants of frontotemporal dementia. Neurology 2004;**62**:742–8.

98 **Rosen HJ**, Narvaez JM, Hallam B, et al. Neuropsychological and functional measures of severity in Alzheimer disease, frontotemporal dementia, and semantic dementia. Alzheimer Dis Assoc Disord 2004;**18**:202–7.

99 **Bier JC**, Ventura M, Donckels V, et al. Is the Addenbrooke's Cognitive Examination effective to detect frontotemporal dementia? J Neurol 2004;**251**:428–31.

100 **Nedjam Z**, Devouche E, Dalla Barba G. Confabulation, but not executive dysfunction discriminate AD from frontotemporal dementia. Eur J Neurol 2004;**11**:728–33.

101 **Osawa A**, Maeshima S, Shimamoto Y, et al. Relationship between cognitive function and regional cerebral blood flow in different types of dementia. Disabil Rehabil 2004;**26**:739–45.

102 **Dudas RB**, Berrios GE, Hodges JR. The Addenbrooke's Cognitive Examination (ACE) in the differential diagnosis of early dementias versus affective disorder. Am J Geriatr Psychiatry 2005;**13**:218–26.

103 **Gold BT**, Balota DA, Cortese MJ, et al. Differing neuropsychological and neuroanatomical correlates of abnormal reading in early-stage semantic dementia and dementia of the Alzheimer type. Neuropsychologia 2005;**43**:833–46.

104 **Lipton AM**, Ohman KA, Womack KB, et al. Subscores of the FAB differentiate frontotemporal lobar degeneration from AD. Neurology 2005;**65**:726–31.

105 **Diehl J**, Monsch AU, Aebi C, et al. Frontotemporal dementia, semantic dementia, and Alzheimer's disease: the contribution of standard neuropsychological tests to differential diagnosis. J Geriatr Psychiatry Neurol 2005;**18**:39–44.

106 **Westmacott R**, Black SE, Freedman M, et al. The contribution of autobiographical significance to semantic memory: evidence from Alzheimer's disease, semantic dementia, and amnesia. Neuropsychologia 2004;**42**:25–48.

107 **Snowden J**, Griffiths H, Neary D. Semantic dementia: Autobiographical contribution to preservation of meaning. Cogn Neuropsychol 1994;**11**:265–88.

108 **Price CC**, Grossman M. Verb agreements during on-line sentence processing in Alzheimer's disease and frontotemporal dementia. Brain Lang 2005;**94**:217–32.

109 **Snowden JS**, Thompson JC, Neary D. Knowledge of famous faces and names in semantic dementia. Brain 2004;**127**:860–72.

110 **Clague F**, Dudas RB, Thompson SA, et al. Multidimensional measures of person knowledge and spatial associative learning: can these be applied to the differentiation of Alzheimer's disease from frontotemporal and vascular dementia? Neuropsychologia 2005;**43**:1338–50.

111 **Fernandez-Duque D**, Black SE. Impaired recognition of negative facial emotions in patients with frontotemporal dementia. Neuropsychologia 2005;**43**:1673–87.

112 **Lee AC**, Rahman S, Hodges JR, et al. Associative and recognition memory for novel objects in dementia: implications for diagnosis. Eur J Neurosci 2003;**18**:1660–70.

113 **Blair M**, Kertesz A, McMonagle P, et al. Quantitative and qualitative analyses of clock drawing in frontotemporal dementia and Alzheimer's disease. J Int Neuropsychol Soc 2006;**12**:159–65.

114 **Cosentino S**, Chute D, Libon D, et al. How does the brain support script comprehension? A study of executive processes and semantic knowledge in dementia. Neuropsychology 2006;**20**:307–18.

115 **Davies RR**, Graham KS, Xuereb JH, et al. The human perirhinal cortex and semantic memory. Eur J Neurosci 2004;**20**:2441–6.

116 **Hou CE**, Miller BL, Kramer JH. Patterns of autobiographical memory loss in dementia. Int J Geriatr Psychiatry 2005;**20**:809–15.

117 **Ivanoiu A**, Cooper JM, Shanks MF, et al. Patterns of impairment in autobiographical memory in the degenerative dementias constrain models of memory. Neuropsychologia 2006;**44**:1936–55.

118 **Mendez MF**, Clark DG, Shapira JS, et al. Speech and language in progressive nonfluent aphasia compared with early Alzheimer's disease. Neurology 2003;**61**:1108–13.

119 **Nestor PJ**, Graham NL, Fryer TD, et al. Progressive non-fluent aphasia is associated with hypometabolism centred on the left anterior insula. Brain 2003;**126**:2406–18.

120 **Scahill VL**, Hodges JR, Graham KS. Can episodic memory tasks differentiate semantic dementia from Alzheimer's disease? Neurocase 2005;**11**:441–51.

121 **Souchay C**, Isingrini M, Pillon B, et al. Metamemory accuracy in Alzheimer's disease and frontotemporal lobe dementia. Neurocase 2003;**9**:482–92.

122 **Vandenberghe RR**, Vandenbulcke M, Weintraub S, et al. Paradoxical features of word finding difficulty in primary progressive aphasia. Ann Neurol 2005;**57**:204–9.

123 **Jenner C**, Reali G, Puopolo M, et al. Can cognitive and behavioural disorders differentiate frontal variant-frontotemporal dementia from Alzheimer's disease at early stages? Behav Neurol 2006;**17**:89–95.

124 **Kertesz A**, Davidson W, McCabe P, et al. Primary progressive aphasia: diagnosis, varieties, evolution. J Int Neuropsychol Soc 2003;**9**:710–19.

125 **Crutch SJ**, Warrington EK. Partial knowledge of abstract words in patients with cortical degenerative conditions. Neuropsychology 2006;**20**:482–9.

126 **Mendez MF**, Anderson E, Shapira JS. An investigation of moral judgement in frontotemporal dementia. Cogn Behav Neurol 2005;**18**:193–7.

127 **Engelborghs S**, Maertens K, Vloeberghs E, et al. Neuropsychological and behavioural correlates of CSF biomarkers in dementia. Neurochem Int 2006;**48**:286–95.

128 **Engelborghs S**, Maertens K, Marien P, et al. Behavioural and neuropsychological correlates of frontal lobe features in dementia. Psychol Med 2006;**36**:1173–82.

129 **Castiglioni S**, Pelati O, Zuffi M, et al. The frontal assessment battery does not differentiate frontotemporal dementia from Alzheimer's disease. Dement Geriatr Cogn Disord 2006;**22**:125–31.

130 **Grossi D**, Fragassi NA, Chiacchio L, et al. Do visuospatial and constructional disturbances differentiate frontal variant of frontotemporal dementia and Alzheimer's disease? An experimental study of a clinical belief. Int J Geriatr Psychiatry 2002;**17**:641–8.

131 **Murtha S**, Cismaru R, Waechter R, et al. Increased variability accompanies frontal lobe damage in dementia. J Int Neuropsychol Soc 2002;**8**:360–72.

132 **Crutch SJ**, Warrington EK. Acalculia: deficits of operational and quantity number knowledge. J Int Neuropsychol Soc 2001;**7**:825–34.

133 **Graham KS**, Simons JS, Pratt KH, *et al.* Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia* 2000;**38**:313–24.

134 **Grossman M**, Mickanin J, Onishi K, *et al.* Progressive nonfluent aphasia: Language, cognitive, and PET measures contrasted with probable Alzheimer's disease. *J Cogn Neurosci* 1996;**8**:135–54.

135 **Westbury C**, Bub D, Chertkow H. Distinct neurolinguistic symptom clusters in Alzheimer's-type dementia and primary progressive aphasia. *Brain Cogn* 2002;**48**:611–17.

136 **Kremin H**, Perrier D, De Wilde M, *et al.* Factors predicting success in picture naming in Alzheimer's disease and primary progressive aphasia. *Brain Cogn* 2001;**46**:180–3.

137 **van Beilen M**, Pijnenborg M, van Zomeren EH, *et al.* What is measured by verbal fluency tests in schizophrenia? *Schizophr Res* 2004;**69**:267–76.

138 **Ardila A**, Pineda D, Rosselli M. Correlation between intelligence test scores and executive function measures. *Arch Clin Neuropsychol* 2000;**15**:31–6.

139 **Demakis GJ**. Frontal lobe damage and tests of executive processing: a meta-analysis of the category test, stroop test, and trail-making test. *J Clin Exp Neuropsychol* 2004;**26**:441–50.

140 **Storey E**, Slavin MJ, Kinsella GJ. Patterns of cognitive impairment in Alzheimer's disease: assessment and differential diagnosis. *Front Biosci* 2002;**7**:e155–84.

141 **Cummings JL**, McPherson S. Neuropsychiatric assessment of Alzheimer's disease and related dementias. *Aging (Milano)* 2001;**13**:240–6.

142 Lichtenberg PA, Murman DL, Mellow AM, eds. *Handbook of dementia:Psychological, neurological, and psychiatric perspectives*. Hoboken, NJ, US: John Wiley and Sons, 2003.

143 **Cummings JL**. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. *Neurology* 1997;**48**(Suppl 6):S10–16.

144 **Hughes CP**, Berg L, Danziger WL, *et al.* A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982;**140**:566–72.

145 **Mychack P**, Rosen H, Miller BL. Novel applications of social-personality measures to the study of dementia. *Neurocase* 2001;**7**:131–43.