

The Differential Effects of Alzheimer's Disease and Lewy Body Pathology on Cognitive Performance: a Meta-analysis

Ashita S. Gurnani¹ · Brandon E. Gavett¹

Received: 24 January 2016 / Accepted: 29 September 2016 / Published online: 22 November 2016
© Springer Science+Business Media New York 2016

Abstract Differential diagnosis of Alzheimer's disease (AD) from normal aging and other dementia etiologies is imperative for disease specific treatment options and long-term care planning. Neuropathological confirmation is the gold standard for neurodegenerative disease diagnosis, yet most published studies examining the use of neuropsychological tests in the differential diagnosis of dementia rely upon clinical diagnostic outcomes. The present study undertook a meta-analytic review of the literature to identify cognitive tests and domains that allow for the differentiation of individuals with AD pathology from individuals with dementia with Lewy Bodies (DLB) pathology and pathology-free individuals. A comprehensive literature search yielded 14 studies that met the inclusion criteria for the present meta-analysis. Six studies comprised 222 decedents with AD compared to 433 normal controls, and eight studies comprised 431 cases of AD compared to 155 decedents with DLB. Results revealed that the effect of having neuropathologically confirmed AD versus DLB lowered performance in the memory domain, and having DLB decreased performance in the visuospatial domain. No single test differed significantly across the AD and DLB groups. For the AD and pathology free comparison, results indicated that that AD was associated with poorer performance on the memory and language domains. With respect

to specific cognitive tests, AD produced lower scores on list learning tests, category fluency, and the Digit Symbol substitution test. The limited number of studies meeting inclusion criteria warrants formulation of guidelines for reporting in clinico-pathological studies; suggested guidelines are provided.

Keywords Alzheimer's disease · Dementia with Lewy bodies · Neuropsychological tests · Autopsy · Neuropathology · Cognitive tests

Due to considerable heterogeneity in the cognitive functioning of older adults with and without dementia (Mungas et al. 2010), the boundaries between cognitive decline associated with normal aging and different dementia etiologies can be blurred (Galvin et al. 2005). Though neurodegenerative disorders, including Alzheimer's disease (AD), dementia with Lewy Bodies (DLB), frontotemporal dementia (FTD), and vascular dementia (VaD) have distinct pathologies, diagnosis in a majority of cases is made on the basis of clinical symptomatology alone (Johns et al. 2009). Unfortunately, clinical diagnosis is sometimes inaccurate compared to neuropathological diagnosis (Dubois et al. 2007; Galasko et al. 1994; Kazee et al. 1993; Varma et al. 1999). Because most research investigating the utility of neuropsychological tests for differential diagnosis of dementia has relied on clinical diagnosis, understanding of the association between neurodegenerative disease and cognitive phenotype is often limited by the inaccuracy of clinical diagnosis. The present study aims to examine differences in cognitive test results associated with different autopsy verified neurodegenerative conditions. Initially, we planned to examine the four most common causes of dementia: AD, DLB, FTD, and VaD, in addition to pathology-free individuals. However, because no studies examining VaD

Electronic supplementary material The online version of this article (doi:10.1007/s11065-016-9334-0) contains supplementary material, which is available to authorized users.

✉ Ashita S. Gurnani
agurnani@uccs.edu

¹ Department of Psychology, University of Colorado at Colorado Springs, 1420 Austin Bluffs Parkway, Colorado Springs, CO 80918, USA

and only one study examining FTD (Elfgren et al. 1994) was identified that met the inclusion criteria used for this study, analysis of these etiologies was not pursued.

Alzheimer's disease (AD) and dementia with Lewy Bodies (DLB) may present with similar clinical manifestations in the early stages of the disease, in which case, distinctions in their neuropsychological profiles may not be apparent until further along in the disease (Hamilton et al. 2008; McKeith et al. 2005; Yoshizawa et al. 2013). Additionally, the core clinical features of DLB that do not typically overlap with AD - such as cognitive fluctuations, visual hallucinations, and parkinsonism - are present at a low rate (McKeith et al. 2005). Thus, with the exception of brain biopsy, there are no definitive biological markers for the *in vivo* diagnosis of AD and DLB, which can make differential diagnosis between neurodegenerative diseases and normal aging challenging (Kraybill et al. 2005; Mollenhauer et al. 2010; Salmon and Bondi 2009; Sloane et al. 2002). Autopsy studies suggest that misdiagnosis occurs in 12–23% of individuals with a clinical diagnosis of AD (Beach et al. 2012; Gaugler et al. 2013). In addition, a recent retrospective analysis of clinical trials suggested that more than half of individuals with a clinical diagnosis of AD had a second form of neuropathology when examined at autopsy (Wang et al. 2012). Therefore, studies reporting on the diagnostic accuracy of clinical tests, when judged against a clinical – as opposed to pathological – criterion standard, are necessarily limited by the accuracy of clinical diagnosis.

The importance of accurate *in vivo* differential diagnosis is perhaps the most apparent in the context of clinical trials, where results can be compromised if participants are misdiagnosed at study entry. Errors in clinical diagnosis could result in the enrollment of individuals with latent neurodegenerative pathology into a putatively healthy control group or inclusion of participants with DLB in an AD trial, for example (Hamilton et al. 2008; McKeith et al. 2005; Salmon and Bondi 2009). Differential diagnosis between DLB and AD is also important for symptom management, since neuroleptics and other pharmacotherapies may be contraindicated for patients with DLB (Galvin et al. 2008; McKeith et al. 2002, 2005).

Some studies examining the use of neuropsychological measures to differentiate between clinically diagnosed cases of AD and DLB have revealed greater impairment on tasks of attention, executive functioning, and visuospatial deficits in DLB compared to AD (Tiraboschi and Guerra 2010). Consistent with these findings, studies have reported that worse performance on the Trail Making Test (TMT) – part A, Rey Auditory Verbal Learning Test (AVLT) word recall, and Benton Visual Form Discrimination is indicative of DLB versus AD (Ferman et al. 2006). Our group has recently identified lower Mini Mental State Examination (MMSE), Logical Memory Immediate, and Logical Memory Delayed scores as

indicative of clinical AD compared to DLB and lower Digit Span Backwards, TMT A and B, and Functional Activities Questionnaire scores as indicative of clinical DLB compared to AD (John et al. *in press*).

Studies comparing neuropathologically confirmed AD and DLB groups have shown relative preservation of verbal memory and worse visuospatial skills in individuals with DLB compared to AD, with tests such as the Rey-Osterrieth Complex Figure Copy, Block Design, and Clock Drawing having value as predictors of DLB (Ferman et al. 2006; Hamilton et al. 2008; Johnson et al. 2005; Yoshizawa et al. 2013). Other studies have found patients with pathologically confirmed AD and DLB to perform similarly on tests of visuospatial ability overall, with impairment in TMT-B and the Wechsler Adult Intelligence Scales – Revised (WAIS-R) Digit Span subtest indicative of DLB and lower scores on confrontation naming indicative of AD (Ferman et al. 2006; Johnson et al. 2005; Kraybill et al. 2005). Overall, findings have been inconsistent when clinically and pathologically diagnosed AD and DLB groups have been compared on neuropsychological outcome measures.

A similar trend of inconsistent findings has been revealed when comparisons are made between clinical AD and normal controls. Tests of verbal memory, naming, and executive functioning have all been described as being the most sensitive for differentiating clinical AD from normal aging (Chen et al. 2001; Elias et al. 2000; Palmer et al. 2003; Salmon 2000; Tierney et al. 1996).

Studies vary in their findings when comparing neuropsychological test performance in pathologically confirmed AD and normal aging, with some studies indicating TMT-A to be a significant predictor of AD (Galvin et al. 2005) and other studies revealing immediate paragraph recall and delayed word list recall as tests that can differentiate definite AD from normal aging (Schmitt et al. 2000). In an example of starkly contrasting findings, Bennett et al. (2006) found no significant differences (correcting for age, sex, and education) between definite AD and controls on a variety of neuropsychological domains other than episodic memory, whereas Johnson and Galvin (2011) found that the definite AD group showed significantly poorer performance on all cognitive tests in comparison to pathology-free controls.

A qualitative review of this literature cannot adequately summarize the data or reconcile methodological differences between studies. Thus, composite effect size measures produced in a meta-analysis can help with reconciling contradictory results by summarizing the effects of AD and DLB pathology on cognitive performance.

Several meta-analytic studies focusing on the differential diagnosis of AD have been reported. Henry et al. (2004) found that clinical AD patients were relatively more impaired on semantic fluency versus phonemic fluency tasks, while

Christensen et al. (1991) reported that both verbal and non-verbal measures were better at discriminating clinical AD from controls. The only meta-analysis that has compared cognitive performance in clinically diagnosed AD and DLB indicated that DLB cases have more difficulties with attention/executive and visuo-perceptual abilities than AD cases (Collerton et al. 2003). It should be noted that these meta-analytic studies relied upon clinical, not pathological, diagnosis, which limits the validity of the conclusions drawn.

The Present Study

The present study uses meta-analysis to estimate the degree to which AD pathology affects neuropsychological test scores compared to decedents with autopsy confirmed DLB and those with autopsy confirmed healthy brains. Although previous evidence is inconsistent regarding the specific cognitive tests that are most affected by AD in comparison to normal aging and DLB, there appears to be some consistency with respect to differential effects of pathology on performance within broader cognitive domains. Thus, for the AD vs. pathology free comparison, it is hypothesized that having definite AD will negatively affect performance in the memory, language, and executive functioning/attention domains. For the AD vs. DLB comparison, it is hypothesized that having definite AD will negatively affect performance in the memory domain whereas having pathologically-confirmed DLB will negatively impact performance in the visuospatial domain. Because pathological outcomes will serve as the criterion standard, the results of this study will avoid the inaccuracies and potential for criterion contamination inherent to clinical diagnosis (Albert et al. 2005; Salmon and Bondi 2009; Walker et al. 2012).

Method and Statistical Analysis

This study was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA; Moher et al. 2009) guidelines.

Literature Search

A comprehensive search of various electronic databases (e.g., PubMed, Scopus, PsycInfo, SUMsearch-2) from January 1990 to May 2013 was conducted to identify all published studies that examined neuropsychological functioning in autopsy-verified cases of AD. The key search terms *Alzheimer's disease*, *Lewy body*, *AD*, and *DLB* were used in combination with *autopsy*, *neuropathology*, *neuropsychology*, and *clinicopathological* to identify relevant articles.

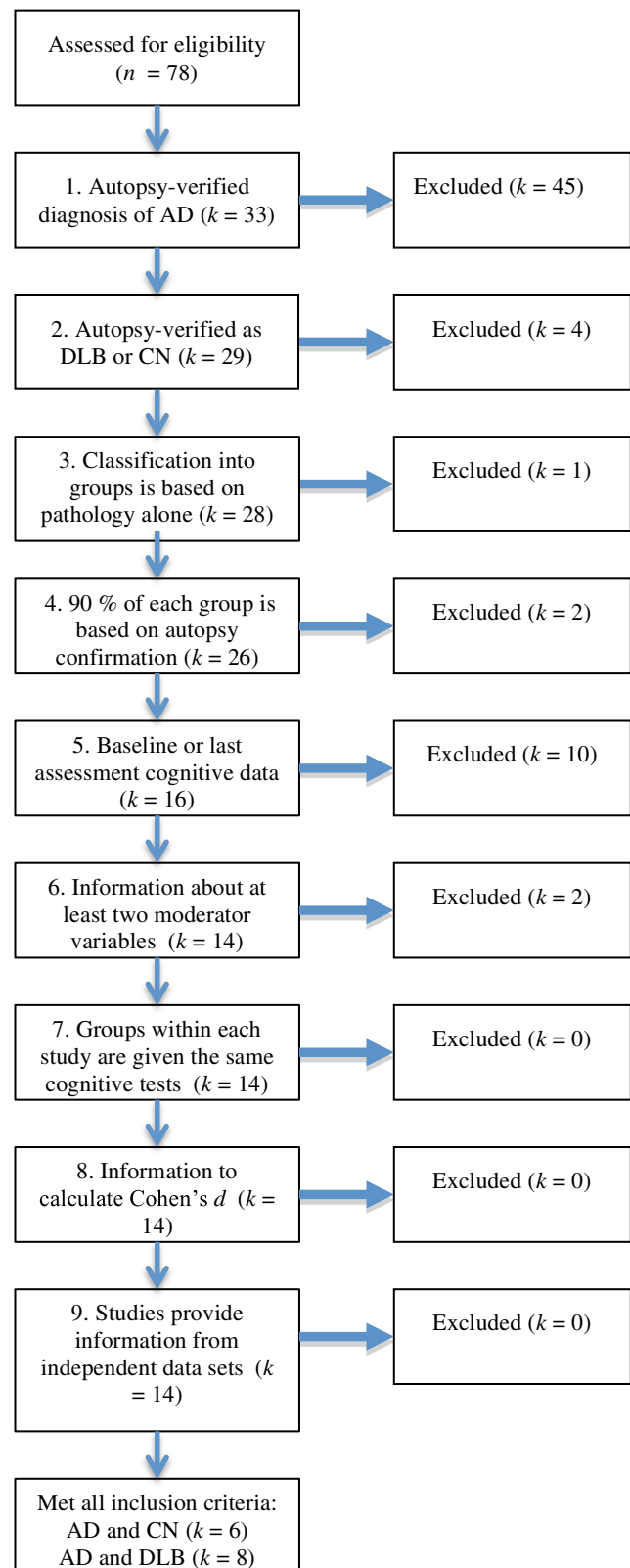


Fig. 1 Flow of article selection through all inclusion criteria in the present meta-analysis. AD = Alzheimer's disease, DLB = dementia with Lewy bodies, CN = Cognitively normal, k = number of studies

Additional references were obtained by examining the bibliographies of relevant papers. Study eligibility was determined by examining the abstract or full text of articles. The search was completed in August 2013 and revealed 78 relevant journal articles (see Fig. 1).

Inclusion Criteria

In order to be selected for the current meta-analysis, a study had to meet the following inclusion criteria. First, the study had to include a participant group that consisted of individuals with an autopsy-verified diagnosis of AD. Out of 78 relevant articles, 33 studies contained participants with autopsy verification of AD. The different neuropathological consensus criteria, including National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA; McKhann et al. 2011) and Khachaturian (Khachaturian 1985) criteria are relatively consistent in their diagnosis of AD (Newell et al. 1999). Thus, studies were included in the meta-analysis irrespective of the criteria used for neuropathological confirmation.

Second, the study had to include a comparison group of individuals that were neuropathologically diagnosed as normal (pathology-free) or with DLB. All studies involving a DLB participant group used the consensus criteria established by McKeith et al. (2005) for neuropathological confirmation. Three out of the 33 studies (Caramelli et al. 1998; Davis et al. 1999; Johnson et al. 2009) did not include a pathologically normal or DLB comparison group and were excluded from the meta-analysis. Although the study conducted by Kanne et al. (1998) included a cognitively normal group, it did not confirm their diagnosis via autopsy, thus, this study was excluded from the present meta-analysis.

Third, participants in each study had to be classified into groups based on pathological confirmation alone irrespective of whether or not they were given a clinical diagnosis in the study. Thus, clinical diagnosis did not inform the classification of participants into AD, DLB or pathology-free groups. Though Galvin et al. (2005) classified their groups based on autopsy confirmation, they presented neuropsychological data on the basis of clinical diagnosis. Additionally, they reported data on some individuals who were included in the study by Johnson et al. (2008). Thus, the Galvin et al. (2005) study was excluded from the meta-analysis.

The remaining 28 studies consisted of participant groups formed on the basis of autopsy verification. The fourth inclusion criterion required at least 90% of all individuals in a group be confirmed as having AD, DLB, or to be free from pathology by autopsy verification. Two out of the remaining 28 studies (Ferman et al. 2006; Johnson et al. 2008) were excluded because they contained participant groups in which

approximately 80% or less of individuals received neuropathological confirmation.

A review of the studies for each comparison group (AD vs. pathologically normal and AD vs. DLB) indicated that most studies in the AD vs. pathologically normal comparison provided data from the last assessment before death, whereas studies comparing AD and DLB provided data taken from a baseline study visit. In an effort to maximize the number of studies within each comparison group, the fifth inclusion criterion required studies comparing AD and pathologically normal groups to provide raw test score data from the last assessment before autopsy whereas studies comparing AD and DLB groups had to provide raw test score data from a baseline visit. Five studies examining AD and pathologically normal groups did not provide last assessment data. Two others (Chui et al. 2006; Reed et al. 2007) provided data for cognitive domains and not for the individual neuropsychological tests. Another study (Price et al. 1998) only provided the level or severity of impairment on each neuropsychological test and not any raw data, and another (Ahmed et al. 2013) provided data for the study visit at which diagnosis of individuals changed from mild cognitive impairment (MCI) to AD. Driscoll et al. (2006) reported the number of individuals that were administered each neuropsychological test but did not report any test scores and did not respond to requests for data.

Sixth, the studies had to provide information regarding at least two of the following moderator variables: mean age at assessment, education, and time interval from assessment to autopsy. Two studies (Hulette et al. 1998; Lopez et al. 1999) did not provide any of this information and did not respond to any requests for data, thus, these studies were excluded from the meta-analysis.

The seventh inclusion criterion required that each participant group in an individual study be given the same battery of neuropsychological tests. Eighth, studies had to provide enough information about cognitive tests such that it would enable the calculation of Cohen's *d* effect sizes. Thus, the included studies provided sample sizes, means, and standard deviations or standard errors for each participant group on each neuropsychological test included in the study. Ninth, studies included for each comparison group (AD vs. control and AD vs. DLB) had to provide information from independent data sets. Thus, there were a total of 14 studies remaining that met all the inclusion criteria. Six studies were included in the AD versus pathologically normal comparison (Table 2) and eight studies were included in the AD versus DLB comparison (Table 3).

Statistical Analyses

Effect Sizes In a meta-analysis, effect sizes from relevant articles are combined to give a weighted mean estimate of the population effect size (Field and Gillett 2010). Though it

would be ideal to use sensitivity and specificity as effect size estimates, only one study provided diagnostic accuracy statistics, Ballard et al. (1999) reported the positive predictive values (PPV) of 0.92 and 0.80 for DLB and AD respectively against their pathological diagnosis. As a result, Cohen's d (Cohen 1988) was chosen as the effect size measure for the present meta-analysis since it could be easily calculated from information provided in all the studies while providing an unbiased effect size estimate across studies of discrepant sample sizes (Field and Gillett 2010). Cohen's d was calculated to measure the extent to which the mean neuropsychological test scores of decedents with AD differed from pathologically confirmed DLB and pathology-free individuals. Each study reported neuropsychological test scores in the form of means and either standard errors or standard deviations. These data were used to calculate effect sizes for each cognitive test used in each individual study. In addition to examining each test separately, we also grouped the tests into cognitive domains based on the primary ability measured. Though tests measure multiple abilities, a determination about primary ability was based on previous literature (Lezak et al. 2012; Strauss et al. 2006).

For the purpose of consistency, effect sizes were calculated by subtracting scores of the pathologically normal or DLB group from the AD group. On most tests, a higher score is indicative of better performance. Thus, a positive effect size indicates that the AD group performed better on a test while a negative effect size indicates that the AD group performed worse on a test. However, on some timed tests, higher scores reflect worse performance. For these tests (e.g. TMT), effect sizes were calculated by subtracting the scores in the AD group from the scores produced by the two comparison groups to facilitate consistent analysis and interpretation of score differences.

R statistical software (version 3.0.0; Development Core Team 2015) package *metafor* (Viechtbauer 2010) was used to calculate pooled effect sizes, standard errors, and 95% confidence intervals for each neuropsychological test and domain. The mean effect size measures the extent to which the two groups differ on a test, while the standard error measures the variability in the effect sizes across different studies (Hutchinson and Mathais 2007). The width of the confidence interval is indicative of the precision of the effect size estimate. When organized by domain, the hierarchical organization of the data (i.e., most studies reported more than one test result) would have violated the assumption of independence if uncorrected. To address this concern, we estimated random effects at the level of the source study in a mixed effects framework, similar to the structure often applied to hierarchical data in mixed effects models (Konstantopoulos 2011; Pinheiro and Bates 2000).

Cohen (1988) suggested small, medium, and large values for the d effect size of 0.2, 0.5, and 0.8 respectively. This effect size is scaled like a z -score; as such, an effect size of 0.5

indicates that the mean scores for the two comparison groups differ by half of a pooled standard deviation (Zakzanis 2001). Various effect size magnitudes have been discussed in the context of group differentiation in psychology. Zakzanis (2001) suggested the use of Cohen's d of 3 as a clinical marker to discriminate almost all individuals across comparison groups. Based on taxometric research in psychopathology, Meehl (1995) suggested the use of Cohen's d of 1.25 and above as a valid marker for diagnostic classification. Hutchinson and Mathais (2007) as well as Mathais and Burke (2009) conducted meta-analyses examining cognitive functioning in clinically diagnosed AD using $d \geq 0.8$ as a marker for deciding whether a cognitive test was useful in differentiating AD from the comparison group.

When it can be assumed that data from the samples being compared have equal variances, equal frequencies, and are normally distributed, Cohen's d can be converted to other statistics, such as area under the receiver operating characteristic curve (AUC; Ruscio 2008). However, both d and AUC simply reflect a test's overall correct classification ability without providing information about rates of false positive and false negative errors. This more detailed information is provided by sensitivity and specificity statistics. Using the same assumptions described above, Cohen's d can be used to approximate a test's optimal sensitivity and specificity when converted to Cohen's U_2 statistic (Cohen 1988). Table 1 can be used to derive Cohen's U_2 and AUC values for a given test based on the Cohen's d statistics derived from this meta-analysis. This table should be used with caution, however, due to the fact that the data used in this meta-analysis are not likely to meet the assumptions discussed above.

Dependent Measures and Model The predictor variable was the diagnostic group (AD, DLB, or pathologically normal), whereas the dependent variable was the effect size estimate derived from the source studies. A random-effects meta-analytic model was used since it was believed that the average effect size of the population varies by study. This model accounts for two error estimates: error created due to sampling of studies from a population of studies and error associated with sampling studies from several populations with different effect sizes (Field and Gillett 2010).

Moderator Analyses The potential influences of age at time of assessment, education, and time interval from assessment to autopsy on cognitive test scores were also examined. The means of the above mentioned moderator variables were calculated for each study by combining data from the AD group and pathologically normal or DLB group, weighted by sample size (Hutchinson and Mathais 2007). In order to examine the relationship between the moderator variables and effect sizes, mixed effects linear modeling

Table 1 Relationship between Cohen's d , Cohen's U_2 , and Area under the curve (AUC)

d	U_2	AUC
0.0	0.5	0.5
0.1	0.52	0.53
0.2	0.54	0.56
0.3	0.56	0.58
0.4	0.58	0.61
0.5	0.6	0.64
0.6	0.62	0.66
0.7	0.64	0.69
0.8	0.66	0.71
0.9	0.67	0.74
1.0	0.69	0.76
1.1	0.71	0.78
1.2	0.73	0.8
1.3	0.74	0.82
1.4	0.76	0.84
1.5	0.77	0.86
1.6	0.79	0.87
1.7	0.8	0.89
1.8	0.82	0.9
1.9	0.83	0.91
2.0	0.84	0.92

AUC = Area under the curve. Cohen's U_2 represents the optimal values for a test's sensitivity and specificity. Deriving AUC and U_2 from Cohen's d require the assumptions of equal variances, equal frequencies, and Gaussian distributions to be met

with random effects was used to account for non-independence of studies within a given domain.

Publication Bias Analyses A fail safe N (Nfs) estimates the number of unpublished studies with null effects that could call into question the findings of the present study. When significant effect size estimates were found, Nfs was calculated in order to identify any bias in the effect size estimates caused by the “file drawer problem” (Hutchinson and Mathais 2007; Rosenthal 1979; Rosenthal and DiMatteo 2001).

Results

The demographic characteristics for participants in both comparisons, AD vs. pathologically normal and AD vs. DLB, are shown in Tables 2 and 3, respectively. Overall, 222 individuals with pathologically confirmed AD and 433 pathologically confirmed normal individuals across six studies were included in the AD versus pathologically normal comparison. Across eight studies, the AD versus DLB comparison included 431

Table 2 Demographic characteristics for AD and pathologically normal groups

Study	AD(N)	PN(N)	Age	Education	Interval
Schmitt et al. (2000)	7	52	83.90	16.30	8.02
Goldman et al. (2001)	10	9	82.25	15.05	9.00
Bennett et al. (2006)	50	84	83.30	17.25	7.08
Johnson et al. (2008)	115	191	76.73	14.40	18.00
Scheff et al. (2011)	8	9	84.80	16.40	7.00
Riley et al. (2011)	32	88	83.40	16.02	8.50

Age and education are presented in number of years. Interval refers to time interval from assessment to autopsy and is presented in number of months. AD, Alzheimer's disease, PN, pathologically normal, N , sample size

individuals with pathologically confirmed AD and 155 individuals with pathologically confirmed DLB.

AD versus Pathologically Normal: Last Assessment before Autopsy

Individual Tests Effect size estimates were calculated for 15 neuropsychological tests that were included in more than one study. Cohen's d estimates, 95% confidence intervals, and Nfs for each of the 15 tests are provided in Table 4. As expected, all effect sizes were negative, which indicates that AD pathology serves to decrease cognitive test performance relative to controls. Results revealed that the effect of having neuropathologically confirmed AD versus no pathology caused significantly lower scores on the following ten cognitive tests: Benton Visual Form Recognition Copy, WAIS Block Design, Boston Naming Test, Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Immediate Recall, CERAD Delay Recall, Category

Table 3 Demographic characteristics for AD and DLB groups

Study	AD(N)	DLB(N)	Age	Education	Interval
Galasko et al. (1996)	38	6	69.61	13.94	100.32
Ballard et al. (1999)	40	40	79.85	–	59.40
Minoshima et al. (2001)	10	4	69.60	–	38.28
Hamilton et al. (2004)	44	22	74.30	14.70	57.60
Kraybill et al. (2005)	48	22	77.15	11.60	70.92
Tiraboschi et al. (2006)	94	23	74.60	14.50	129.60
Beach et al. (2009)	68	26	82.22	–	86.28
Yoshizawa et al. (2013)	89	12	71.60	15.02	124.80

Age and education are presented in number of years; time interval from assessment to autopsy is presented in number of months. AD, Alzheimer's disease; DLB, dementia with Lewy bodies; N = sample size

Table 4 Weighted effect size estimates for cognitive tests comparing AD to pathologically normal groups

Test	<i>k</i>	<i>d</i>	Lower 95% CI	Upper 95% CI	<i>Nfs</i>	<i>p</i>
CERAD Immediate Recall	3	-1.95	-3.01	-0.90	11	< .01
Digit Symbol Substitution Test	2	-1.76	-2.27	-1.24	93	< .01
WMS Associate Learning	2	-1.22	-2.51	0.08	–	.07
WAIS Block Design	2	-1.17	-1.64	-0.70	47	< .01
Boston Naming Test	3	-1.04	-1.96	-0.13	99	.03
WAIS Information Subtest	2	-1.03	-2.02	-0.04	51	.04
WMS Mental Control	2	-0.99	-1.23	-0.76	34	< .01
Trail Making - Part A	3	-0.91	-2.06	0.24	–	.12
Benton VFR Copy	2	-0.72	-1.31	-0.13	21	.02
Category Fluency	4	-0.64	-1.07	-0.21	69	< .01
Logical Memory	5	-0.63	-1.57	0.30	–	.18
MMSE	4	-0.62	-1.30	0.05	–	.07
CERAD Delayed Recall	3	-0.61	-0.87	-0.35	20	< .01
Digit Span Backward	3	-0.58	-0.94	-0.22	30	< .01
Digit Span Forward	3	-0.31	-0.63	0.00	–	.05

Data obtained at the last assessment before autopsy. *k* = number of included studies; *d* = Cohen's *d*; *CI*, confidence interval; *Nfs* = fail safe *N*; *CERAD*, consortium to establish a registry for Alzheimer's Disease; *WMS*, Wechsler Memory Scale; *WAIS*, wechsler Adult intelligence scale; *VFR*, visual form recognition; *MMSE*, mini-mental state examination

Fluency, Digit Span Backward, Digit Symbol substitution test (DSST), WMS Mental Control, and WAIS Information subtest.

Cognitive Domains Effect size estimates were calculated for the domains of attention, processing speed, and executive functioning, language, mental status, memory, and visuospatial ability. The cognitive tests used for each domain are available in Online Resource 1. Results revealed that having AD versus no pathology decreased performance on the language ($d = -0.72$, 95% CI $[-1.32, -0.12]$, $p < .01$, $Nfs = 423$) and memory ($d = -0.79$, 95% CI $[-1.37, -0.21]$, $p = .01$, $Nfs = 1114$) domains. The mental status domain was associated with the largest effect size but the most imprecise confidence intervals ($d = -1.01$, 95% CI $[-2.53, 0.51]$, $p = .20$), while the attention, processing speed, and executive functioning domain had the smallest effect size, but the most precise confidence intervals ($d = -0.48$, 95% CI $[-0.99, 0.03]$, $p = .06$). The visuospatial domain was comparable to the attention, processing speed, and executive functioning domain, both in terms of effect size and confidence intervals ($d = -0.51$, 95% CI $[-1.04, 0.01]$, $p = .06$). Forest plots showing the AD versus pathologically normal comparisons across each domain are presented in Fig. 2.

Moderator Analysis Tests from all cognitive domains were combined for the moderator analysis. Results revealed that the effect of AD pathology on cognitive test results, compared to controls, was not significantly affected by age, $b = -0.80$, 95% CI $[-1.64, 0.32]$, $p = .36$, education, $b = 0.04$, 95% CI

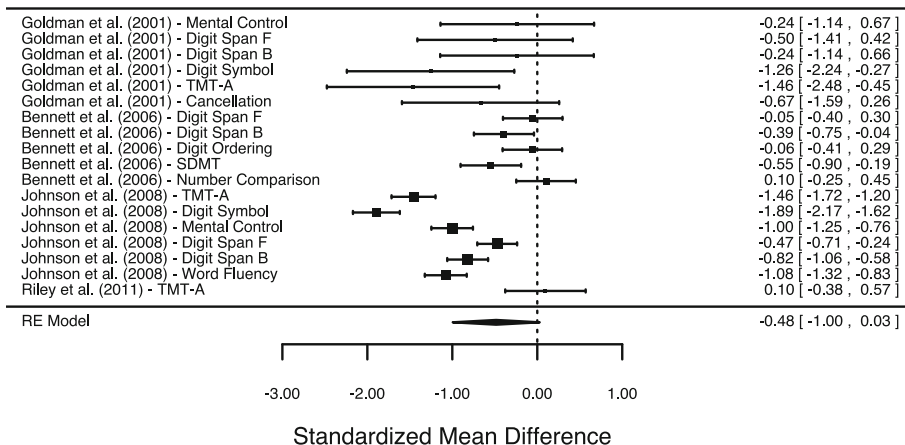
$[-1.53, 1.62]$, $p = .96$, or time interval from assessment to autopsy, $b = -0.56$, 95% CI $[-1.79, 0.70]$, $p = .37$.

AD versus DLB: Baseline Assessment

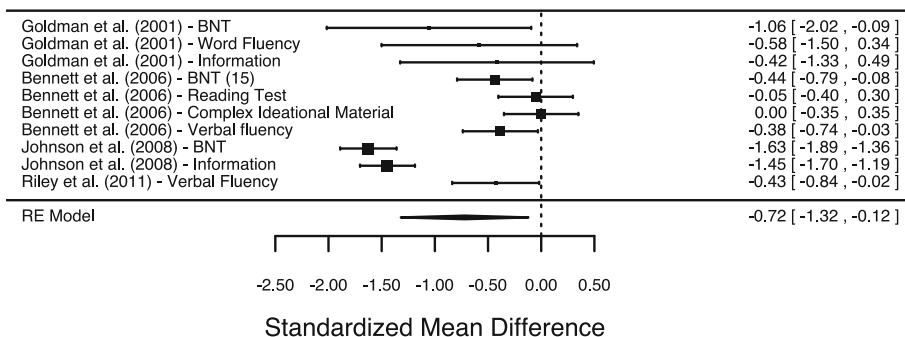
Individual Tests Effect size estimates were calculated for the following four neuropsychological tests that were included in more than one study: WAIS Block Design, Clinical Dementia Rating (CDR), Dementia Rating Scale (DRS), and MMSE. Overall, the effect size estimates ranged from a minimum of -0.15 (95% CI $[-0.37, 0.07]$, $p = .18$) for the MMSE to a maximum of 0.31 (95% CI $[-0.06, 0.67]$, $p = .10$) for the WAIS Block Design test. The effect size estimates for the DRS and CDR were 0.22 (95% CI $[-0.07, 0.50]$, $p = .14$) and 0.30 (95% CI $[-0.23, 0.84]$, $p = .27$), respectively. These results suggest that having DLB versus AD did not significantly affect performance on any of the four neuropsychological tests used in more than one study.

Cognitive Domains Effect size estimates were calculated for the domains of attention, processing speed, and executive functioning, language, mental status, memory, and visuospatial ability. The cognitive tests used for each domain are available in Online Resource 1. Results revealed that having AD, compared to DLB, was associated with better performance in the visuospatial domain ($d = 0.33$, 95% CI $[0.05, 0.60]$, $p = .02$, $Nfs = 5$) and worse performance in the memory domain ($d = -0.53$, 95% CI $[-0.74, -0.32]$, $p < .01$, $Nfs = 209$). In comparison, effect size estimates for the language ($d = -0.24$, 95% CI $[-0.60, 0.12]$, $p = .19$), mental status ($d = -0.02$, 95% CI $[-0.20, 0.15]$, $p = .82$), and attention,

AD vs. Control: Last Assessment Attention, Processing Speed, & Executive Functioning



Language



Memory

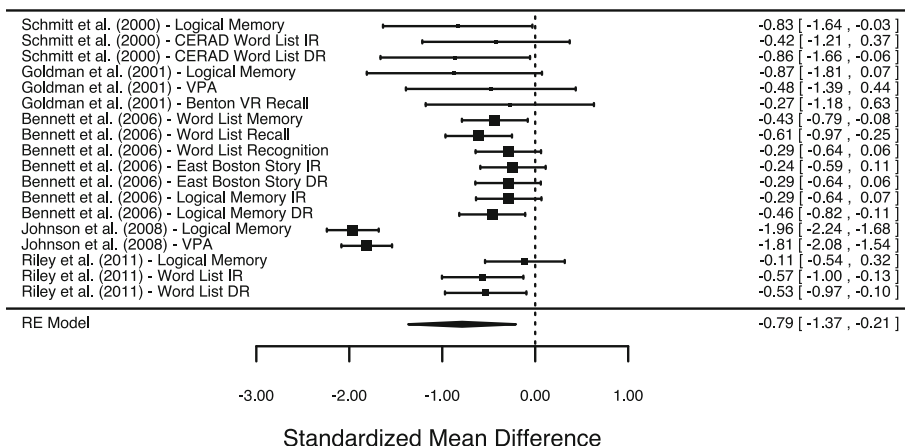


Fig. 2 Forest plot showing the effect size estimates for pathologically confirmed AD compared to pathology-free individuals across five cognitive domains. Data were obtained from the last assessment prior to autopsy. Point size is used to denote relative weightings of each study in the

models. The row labeled RE Model shows the pooled effect size and associated 95% confidence intervals from the random effects meta-analysis, accounting for the hierarchical organization of the data (test nested within study)

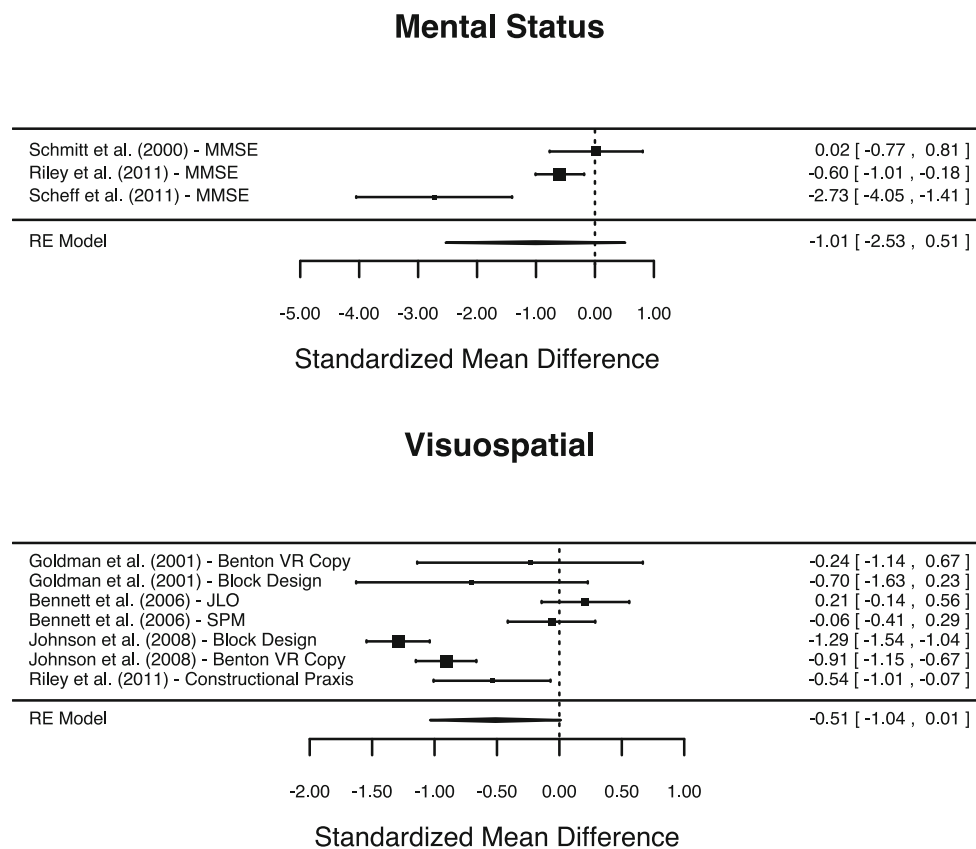


Fig. 2 (continued)

processing speed, and executive function ($d = 0.07$, 95% CI [-0.45, 0.60], $p = .78$), domains were smaller with confidence intervals that spanned zero. Forest plots showing the AD versus DLB comparisons across each domain are presented in Figure 3.

Moderator Analysis Tests from all cognitive domains were combined for the moderator analysis. Results revealed that the differential effects of AD and DLB pathologies on cognitive test results were not significantly affected by age, $b = 0.15$, 95% CI [-0.03, 0.33], $p = .12$, education, $b = 0.20$, 95% CI [-0.10, 0.49], $p = .20$, or time interval from assessment to autopsy, $b = -0.02$, 95% CI [-0.14, 0.11], $p = .79$.

Discussion

This meta-analysis systematically reviewed and quantified the effects of Alzheimer's disease and Lewy body pathology on cognitive performance. Results from the present study indicate that in comparison to pathology free individuals, AD pathology was associated with poorer performance on the DSST, category fluency, and list learning tests at the last assessment before autopsy. With respect to cognitive domains, having AD lowered performance on the memory and

language domains. There was no significant difference in performance on individual cognitive tests between the AD and DLB groups. However, having neuropathologically-confirmed DLB versus AD lowered performance in the visuospatial domain, and having AD pathology decreased performance in the memory domain at the baseline visit. Age, education, and time interval from assessment to autopsy did not moderate the effects of pathology on cognitive performance in either of the comparisons.

AD versus Pathologically Normal: Last Assessment before Autopsy

Individual Tests Cohen's d estimates were obtained for 15 neuropsychological tests, of which 10 tests produced d values with confidence intervals that did not span zero (Table 4). Overall, the Nfs ranged from 11 to 99 for CERAD Immediate Recall and the Boston Naming Test, respectively, making it unlikely that failure to account for unpublished studies could bias the current findings. The DSST ($d = -1.76$) and CERAD Immediate Recall ($d = -1.95$) had the largest effect sizes; these values of d are indicative of optimal sensitivity (and specificity) values of .81 and .84, respectively, given the assumptions discussed previously. However, it is important to note that the confidence intervals

AD vs. DLB: Baseline Assessment Attention, Processing Speed, & Executive Functioning

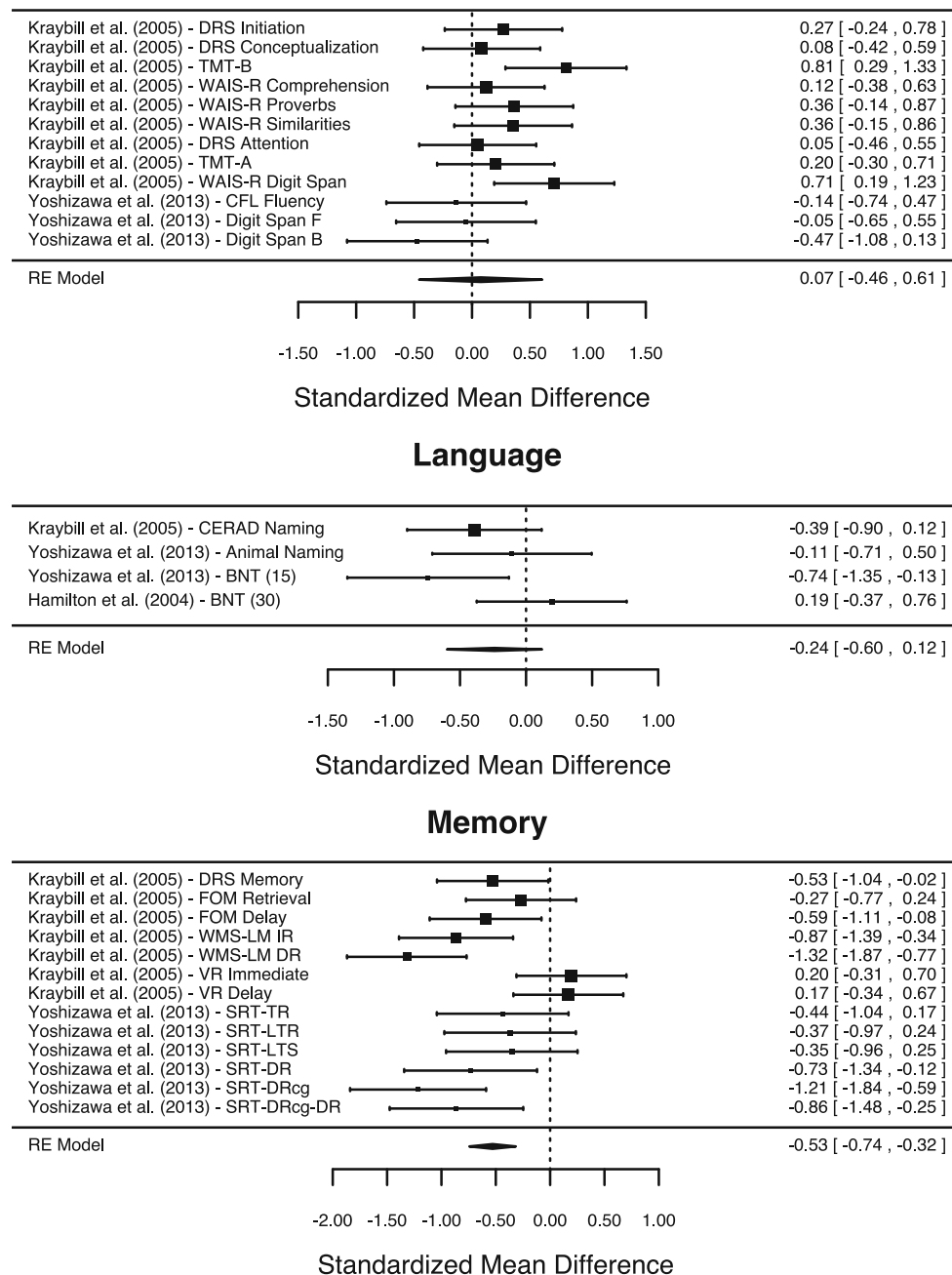


Fig. 3 Forest plot showing the effect size estimates for comparisons between pathologically confirmed AD and DLB across five cognitive domains. Data were obtained from the baseline assessment. Point size is used to denote relative weightings of each study. The row labeled RE

Model shows the pooled effect size and associated 95% confidence intervals from the random effects meta-analysis, accounting for the hierarchical organization of the data (test nested within study)

for both tests were relatively imprecise due to the inclusion of only a few studies. Poor performance on the DSST has been found to be associated with greater burden of neurofibrillary tangles in the limbic region of the brain, suggesting that performance on this test is sensitive to AD pathology (Price et al. 2009; Salmon and Bondi 2009; Tabert et al. 2006). Overall,

the DSST appears to have better diagnostic utility than other tests of attention and psychomotor speed such as Digit Span and Trail Making.

Studies using both CERAD Immediate and Delayed Recall tests have shown that individuals with pathologically and clinically-diagnosed AD obtain significantly lower scores on

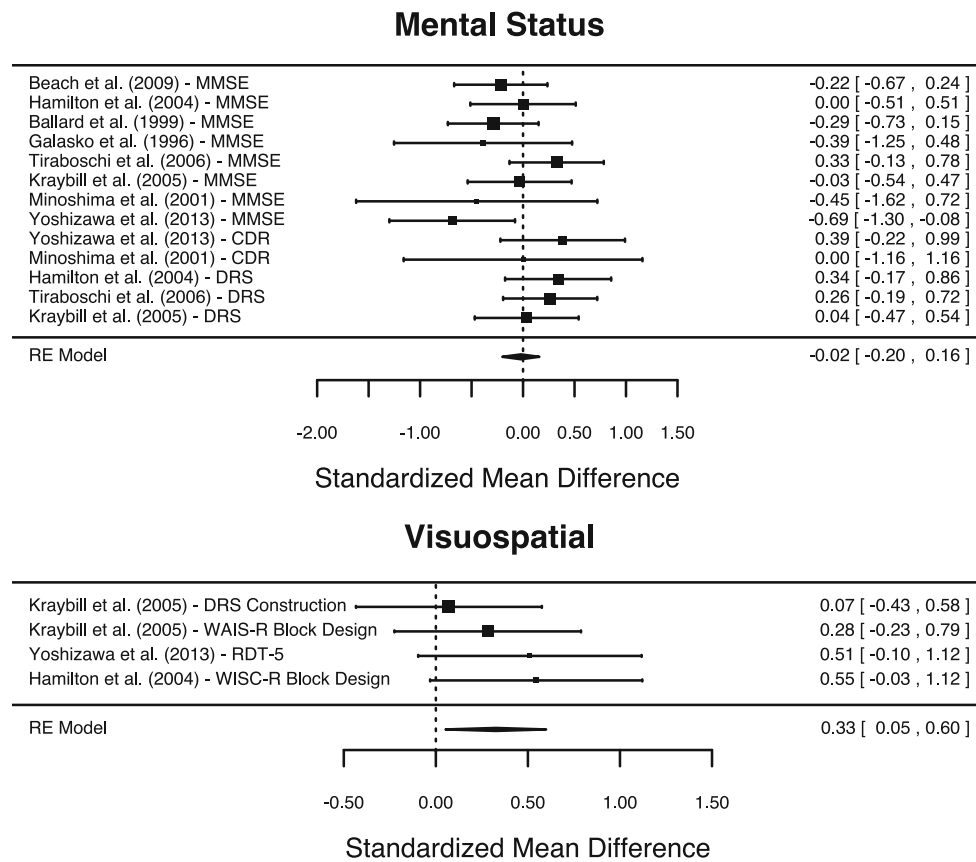


Fig. 3 (continued)

the delayed recall trial in comparison to healthy controls (Schmitt et al. 2000; Welsh et al. 1991). Delayed recall on list learning tests similar to the CERAD's, such as the AVLT, has been shown to predict the clinical diagnosis of AD with a high degree of accuracy (Tierney et al. 1996). In the present study, CERAD Immediate Recall has a wider confidence interval $[-3.01, -0.90]$ but a larger effect size estimate ($d = -1.95$), whereas CERAD Delayed Recall has a narrower confidence interval $[-0.87, -0.35]$ but a smaller effect size estimate ($d = -0.61$). However, because the two confidence intervals do not overlap, these results suggest that CERAD Immediate Recall was affected to a greater degree than Delayed Recall when comparing definite AD to pathology free controls. The results from CERAD Immediate and Delayed Recall together confirm previous evidence that acquisition and recall of verbal information learned over consecutive trials is possibly one of the most valid predictors of AD (Albert et al. 2001; Jacobs et al. 1995; Tabert et al. 2006).

WMS Mental Control had a large effect size estimate ($d = -0.99$, 95% CI $[-1.23, -0.76]$), with the most precise confidence intervals of the 15 cognitive tests. This finding can be interpreted to suggest that, under certain assumptions (discussed above), Mental Control may possess optimal sensitivity and specificity values of 0.69 for differentiating between AD and controls. The confidence intervals obtained

were as narrow as those reported in other studies making similar comparisons (Goldman et al. 2001; Johnson et al. 2008; Tierney et al. 1987). The Mental Control test measures attention and working memory and has been found to be useful in distinguishing healthy controls from those with clinically diagnosed AD (Storandt et al. 1984; Tierney et al. 1987, 1996). Though deficits in attention abilities typically follow memory decline in AD, performance on tasks such as Mental Control that require sustained attention and working memory are also impaired earlier on in the disease process (Lamar et al. 2002; Salmon and Bondi 2009).

Cognitive Domains Results from cognitive domain analysis supported the hypothesis that, in comparison to pathology-free individuals, AD pathology lowered performance on memory and language domains. The biggest effect size was obtained for the mental status domain ($d = -1.01$, 95% CI $[-2.53, 0.51]$), though the wide confidence intervals reveal that the effect of AD pathology on this domain could not be estimated precisely. This is roughly equivalent to optimal sensitivity and/or specificity values of 0.69. The imprecision of this estimate is likely because this effect size was only estimated by three studies, one of which had a total sample size of 17 (Scheff et al. 2011). The only test included in the mental status domain was the MMSE, which has consistently been

shown to be useful in tracking cognitive impairment over time, but not as a diagnostic tool for the detection of dementia (Mitchell 2009; Tombaugh and McIntyre 1992).

Consistent with the results based on individual cognitive tests and in comparison with other domains evaluated, the domain of memory had a large effect size ($d = -0.79$, 95% CI $[-1.36, -0.21]$) with the narrowest confidence intervals. This effect size – roughly equivalent to optimal sensitivity and specificity values of 0.65 – was smaller than the one obtained in a meta-analysis conducted by Bäckman et al. (2005). One explanation for this discrepancy could be that Bäckman et al. (2005) divided their memory domain into primary and episodic memory, thereby increasing the specificity of the memory abilities assessed. In addition, Bäckman's increased precision is likely the result of a larger k , which is a consequence of using a clinical – as opposed to pathological – criterion standard.

Similar to the memory domain, the language domain also had a large effect size ($d = -0.72$, 95% CI $[-1.32, -0.12]$), but with a wider confidence interval. As shown in Table 1, a Cohen's d value of .72 reflects optimal sensitivity/specificity values of 0.64 when the assumptions discussed above are met. The language domain included tests measuring confrontation naming, semantic memory, fluency, comprehension, and reading, as described above, individual tests of confrontation naming, semantic memory, and category fluency were differentially affected by AD pathology. The results from the cognitive domain analyses are consistent with research indicating that tests of episodic and semantic memory both tend to be affected – but to different degrees – by AD pathology (e.g., Hirni et al. 2016).

Moderator analyses revealed that time interval from assessment to autopsy, age, and education did not moderate the effect of AD pathology on cognitive test performance. However, these findings should not be considered conclusive at this point, as only a small subset of ages, education levels, and intervals were available for analysis in the current study (see Table 3).

AD versus DLB: Baseline Assessment

Individual Tests Cohen's d estimates obtained for the four individual neuropsychological tests, including Block Design, CDR, DRS, and MMSE were small with wide confidence intervals that spanned zero. The width of the confidence intervals indicates that findings based on single cognitive tests are extremely imprecise and therefore of limited utility. Previous evidence suggests that, in comparison to those with clinically-diagnosed AD, individuals with clinically-diagnosed DLB perform better on a variety of cognitive measures including Trail Making A, Boston Naming Test, and list learning tests, but worse on tests examining visuospatial ability such as Block Design (Ferman et al. 2006; Hamilton et al. 2008; Kraybill et al. 2005). Studies have also shown that individuals with clinically-diagnosed DLB exhibit deficits in attention and executive function as evidenced by lower scores on digit span and

the attention domain of the DRS in comparison to those with clinically-diagnosed AD (Salmon and Bondi 2009). These discrepancies highlight the need for more studies examining cognitive test performance in pathologically confirmed AD and DLB, which can provide more information about the tests that are most affected by AD or DLB pathology and can also help improve the precision of the effect size estimates.

Cognitive Domains Of the five cognitive domains analyzed in the present study, the domains of memory and visuospatial ability were associated with significant differences. Though moderate in size, the largest effect size was obtained for the memory domain ($d = -0.53$, 95% CI $[-0.74, -0.32]$) with relatively narrow confidence intervals, suggesting that in comparison to DLB, AD pathology negatively impacts performance on tests assessing memory. This effect size reflects optimal sensitivity/specificity values of 0.60 according to Cohen's U_2 . Studies have shown that in comparison to clinically and pathologically diagnosed AD, verbal memory and episodic memory are relatively preserved in clinically and pathologically diagnosed DLB (Hamilton et al. 2004; Johnson et al. 2005). The moderate effect size obtained in the present study was based on tests examining different aspects of memory, including visual and verbal memory. For example, the memory domain included tests of visual memory such as Visual Reproduction, tests of verbal memory such as Logical Memory and the Selective Reminding Test, and tests requiring both verbal and visual memory skills, such as the Fuld Object Memory Evaluation. The inclusion of visually based tests in the memory domain may bias this effect size estimate, given that DLB pathology is associated with lower visuospatial abilities than AD pathology, as discussed next.

The hypothesis that DLB pathology would differentially affect performance on the visuospatial domain in comparison to AD pathology was supported. The effect size estimate for the visuospatial domain ($d = 0.33$, 95% CI $[0.05, 0.60]$) was small in size with relatively narrow confidence intervals. Thus, DLB pathology appears to cause a moderate but detectable reduction in visuospatial skills relative to AD. This result is consistent with previous research that has shown poor performance on the visuospatial domain to be a distinguishing factor between AD and DLB, with clinical and autopsy studies indicating that individuals with DLB perform worse on tests that involve visuoconstruction and visuo-perceptual abilities (Ferman et al. 2006; Hamilton et al. 2008; Johnson et al. 2005; Tiraboschi et al. 2006). However, the estimated effect of DLB, compared to AD, on the visuospatial domain was small, suggesting a good deal of overlap between the two dementia etiologies on visuospatial test scores. This effect size estimate is indicative of optimal sensitivity and specificity of only 0.57 under the assumptions discussed earlier in the manuscript. Therefore, visuospatial tests may not differentiate between AD and DLB as accurately as clinical studies – or

clinical lore – may suggest. In clinical settings, the expectation that visuospatial skills accurately distinguish DLB from AD may serve as a self-fulfilling prophecy, whereby AD patients with visuospatial deficits are mistakenly diagnosed as having DLB, and vice versa. Our reliance on pathological confirmation as our criterion standard avoids the potential for such criterion contamination, though it bears mention that the pathological overlap between AD and DLB is fairly common (Josephs et al. 2004; Mrazek and Griffin 2007).

Similarly, the attention and executive deficits previously reported as capable of differentiating DLB from AD were not supported ($d = 0.07$, 95% CI $[-0.45, 0.60]$). The results of the present study revealed a small effect size estimate with relatively wide confidence intervals that spanned zero. This effect size estimate indicates optimal sensitivity and specificity values of just 0.51. Thus, results from the present study differ from previous findings in the literature that suggest attention and executive dysfunction can help with the differential diagnosis between AD and DLB. On one hand, past research about the cognitive profiles of AD and DLB are primarily based on clinical diagnosis, which introduces potential criterion contamination and diagnostic inaccuracies. On the other hand, the results from this meta-analysis are imprecise due to the inclusion of a small number of studies. The discrepancy between the results from the present study and the literature can be clarified by conducting additional clinicopathological studies that examine specific cognitive abilities in AD and DLB. Finally, moderator analyses revealed that time interval from assessment to autopsy, age, and education did not influence cognitive test performance.

Limitations and Future Directions

Although the present meta-analysis revealed some large and medium effect sizes that were estimated with relatively precise confidence intervals, many other estimates were imprecise. It is likely that the small number of studies that met our inclusion criteria limited the precision of the effect size estimates, thereby decreasing the utility of the results. Many tests that are routinely administered in clinical and research settings were not available for inclusion in the present study. Since differential diagnosis relies heavily on the assumption that disease pathology is affecting task performance, there is a need for many more clinicopathological studies examining cognitive performance across various neurodegenerative conditions. Importantly, studies are inconsistent in the data that they provide, which hinders the process of meta-analysis and establishment of cumulative scientific knowledge. Published clinicopathological studies differ with respect to the cognitive tests used, evaluation time points, reporting of adjusted versus unadjusted test scores, and information reported about moderator variables, which limits their

potential for inclusion in meta-analyses. To facilitate the application of meta-analysis across clinicopathological neurodegenerative disease research, we propose the following reporting guidelines that are recommended for future studies in this area.

Clinicopathological Reporting Guidelines in Neurodegenerative Disease Research

General Guidelines

1. All of the reporting guidelines below should be applied separately to different pathologically verified groups (e.g., AD, DLB, pathologically normal).
2. Report whether the data are original or whether they have been previously published in other studies.
3. Provide a detailed report of participant recruitment, as well as inclusion and exclusion criteria; indicate whether excluded participants differed from included participants in relevant areas.
4. Studies of diagnostic accuracy should report information in accordance with the guidelines proposed in the Standards for Reporting of Diagnostic Accuracy (STARD; Bossuyt et al. 2003), and should, when possible, report sensitivity and specificity data.

Reporting of Clinical Data

5. Collect and report clinical data (e.g., the same battery of neuropsychological tests) consistently across clinical groups.
6. Report raw test scores for cognitive variables, not standardized or norm-referenced scores.
7. Report effect sizes for all cognitive tests, when possible. For example, when comparing two groups, Cohen's d should be presented, along with its 95% confidence intervals. At a minimum, researchers should always provide sufficient data (i.e., means, standard deviations, and sample sizes) to allow other researchers to calculate effect sizes, even if effect sizes are also presented.
8. Report all relevant demographic variables, including age at assessment, age at death, years of education, sex, race, ethnicity, handedness, language, and so forth.
9. Report the length of the time interval between clinical assessment and death.
10. For longitudinal studies, report whether the clinical data were obtained from the first (baseline) study visit, the last study visit before death, or some other time point.
11. For longitudinal studies, report the number of previous study visits the participants had completed prior to the assessment from which data were derived.
12. Report the clinical diagnosis of participants that was made during the reported study visit.

13. Report demographic and genetic differences between decedents and living study participants, to evaluate selection bias.

Reporting of Pathology Data

14. Report the number and percentage of decedents who were studied at autopsy, as a portion of the entire sample used to derive the data being reported.
15. Report the neuropathological criteria used to diagnose decedents (e.g., NIA-Reagan).
16. Report the methods for sampling and quantifying brain tissue pathology. When possible, follow the guidelines put forth by Hyman et al. (2012).
17. Report appropriate genetic variables (e.g., APOE genotype) and the method used to analyze genetic data.
18. If a cognitively normal comparison group was used, report whether this group was verified as being free from pathology at autopsy.

Conclusions

The current study used meta-analysis to systematically review and quantitatively aggregate cognitive test scores to determine the individual tests and cognitive domains that are differentially affected by AD compared to both healthy brains and pathologically confirmed DLB. The individual tests that were most affected by AD pathology relative to controls were CERAD Immediate and Delayed Recall, DSST, WAIS Block Design and Information subtests, Boston Naming Test, WMS Mental Control, Benton Visual Form Recognition Copy, Category Fluency, and the Backward Digit Span subtest. The domains of memory and language were most reliably affected by AD pathology relative to controls. Although no individual tests were found to be differentially affected by AD and DLB pathology, the broader domains of memory and visuospatial functioning were differentially affected by the two neuropathological entities. The effect of having AD pathology relative to either DLB pathology or no pathology was not significantly moderated by age, education, or the time interval between assessment and autopsy. The most obvious strength and unique feature of this study is its use of pathological verification as the criterion standard for defining AD, DLB, and control groups. By using pathologically verified samples, we avoided the confounds and inaccuracies inherent to clinical diagnosis and the possibility of criterion contamination. However, pathological verification also limited the number of studies eligible for inclusion, which in turn limited the precision of our estimated effect sizes. Pathological studies also suffer from a lack of consistency in

reporting, which we attempted to address by introducing guidelines for reporting in clinicopathological studies of neurodegenerative diseases. It is our hope that future research will utilize these guidelines to facilitate a better appreciation of the diagnostic validity of clinical assessment instruments.

Acknowledgments The data from this manuscript were presented at the annual conferences of the International Neuropsychological Society (2014, 2015) and the American Academy of Clinical Neuropsychology (2014). This manuscript was based on the first author's Master's thesis.

Compliance with Ethical Standards

Funding None.

Conflict of Interest The authors declare that they have no conflict of interest.

References

References marked with an asterisk indicate studies included in the meta-analysis

- Ahmed, S., de Jager, C. A., Haigh, A. M., & Garrard, P. (2013). Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology*. doi:10.1037/a0031288.
- Albert, M. S., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of the International Neuropsychological Society*. doi:10.1017/S1355617701755105.
- Albert, M. S., DeCarli, C., DeKosky, S., De Leon, M., Foster, N. L., Fox, ... Khachaturian, Z. (2005). The use of MRI and PET for clinical diagnosis of dementia and investigation of cognitive impairment: A consensus report. In Alzheimer's Association.
- Bäckman, L., Jones, S., Berger, A. K., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. *Neuropsychology*. doi:10.1037/0894-4105.19.4.520.
- * Ballard, C., Holmes, C., McKeith, I., Neill, D., O'Brien, J., Cairns, N., et al. (1999). Psychiatric morbidity in dementia with Lewy bodies: a prospective clinical and neuropathological comparative study with Alzheimer's disease. *American Journal of Psychiatry*, 156, 1039–1045.
- * Beach, T. G., Adler, C. H., Lue, L., Sue, L. I., Bachalakuri, J., Henry-Watson, J., et al. (2009). Unified staging system for Lewy body disorders: correlation with nigrostriatal degeneration, cognitive impairment and motor dysfunction. *Acta Neuropathologica*. doi:10.1007/s00401-009-0538.
- Beach, T. G., Monsell, S. E., Phillips, L. E., & Kukull, W. (2012). Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer disease centers, 2005–2010. *Journal of Neuropathology & Experimental Neurology*. doi:10.1097/NEN.0b013e31824b211b.
- * Bennett, D. A., Schneider, J. A., Arvanitakis, Z., Kelly, J. F., Aggarwal, N. T., Shah, R. C., & Wilson, R. S. (2006). Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*. doi:10.1212/01.wnl.0000219668.47116.e6.

- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., et al. (2003). The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Annals of Internal Medicine*, 138, W1–12. doi:10.7326/0003-4819-138-1-200301070-00012-w1.
- Caramelli, P., Robitaille, Y., LarocheCholette, A., Nitrini, R., Gauvreau, D., Joannette, Y., & Lecours, A. R. (1998). Structural correlates of cognitive deficits in a selected group of patients with Alzheimer's disease. *Cognitive and Behavioral Neurology*, 11, 184–190.
- Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S. T., & Ganguli, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. *Archives of General Psychiatry*. doi:10.1001/archpsyc.58.9.853.
- Christensen, H., Hadzi-Pavlovic, D., & Jacomb, P. (1991). *The psychometric differentiation of dementia from normal aging: A meta-analysis*. Psychological Assessment: A Journal of Consulting and Clinical Psychology. doi:10.1037//1040-3590.3.2.147.
- Chui, H. C., Zarow, C., Mack, W. J., Ellis, W. G., Zheng, L., Jagust, W. J., et al. (2006). Cognitive impact of subcortical vascular and Alzheimer's disease pathology. *Annals of Neurology*. doi:10.1002/ana.21009.
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Collerton, D., Burn, D., McKeith, I., & O'Brien, J. (2003). Systematic review and meta-analysis show that dementia with Lewy bodies is a visual-perceptual and attentional-executive dementia. *Dementia and Geriatric Cognitive Disorders*. doi:10.1159/000072807.
- Davis, D. G., Schmitt, F. A., Wekstein, D. R., & Markesbery, W. R. (1999). Alzheimer neuropathologic alterations in aged cognitively normal subjects. *Journal of Neuropathology and Experimental Neurology*. doi:10.1097/00005072-199904000-00008.
- Development Core Team, R. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available from <http://www.R-project.org>.
- Driscoll, I., Resnick, S. M., Troncoso, J. C., An, Y., O'Brien, R., & Zonderman, A. B. (2006). Impact of Alzheimer's pathology on cognitive trajectories in nondemented elderly. *Annals of Neurology*. doi:10.1002/ana.21031.
- Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *The Lancet Neurology*. doi:10.1016/S1474-4422(07)70178-3.
- Elfgren, C., Brun, A., Gustafson, L., Johanson, A., Minthon, L., Passant, U., & Risberg, J. (1994). Neuropsychological tests as discriminators between dementia of Alzheimer type and frontotemporal dementia. *International Journal of Geriatric Psychiatry*. doi:10.1002/gps.930090807.
- Elias, M., Beiser, A., Wolf, P., Au, R., White, R., & D'Agostino, R. (2000). The preclinical phase of Alzheimer disease: a 22-year prospective study of the Framingham cohort. *Archives of Neurology*. doi:10.1001/archneur.57.6.808.
- Ferman, T. J., Smith, G. E., Boeve, B. F., Graff-Radford, N. R., Lucas, J. A., Knopman, D. S., et al. (2006). Neuropsychological differentiation of dementia with Lewy bodies from normal aging and Alzheimer's disease. *The Clinical Neuropsychologist*. doi:10.1080/13854040500376831.
- Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*. doi:10.1348/000711010X502733.
- Galasko, D., Hansen, L. A., Katzman, R., Wiederholt, W., Masliah, E., Terry, R., et al. (1994). Clinical-neuropathological correlations in Alzheimer's disease and related dementias. *Archives of Neurology*. doi:10.1001/archneur.1994.00540210060013.
- * Galasko, D., Katzman, R., Salmon, D. P., & Hansen, L. (1996). Clinical and neuropathological findings in Lewy body dementias. *Brain and Cognition*. doi:10.1006/brcg.1996.0040.
- Galvin, J. E., Powlishta, K. K., Wilkins, K., McKeel, D. W., Xiong, C., Grant, E., et al. (2005). Predictors of preclinical Alzheimer disease and dementia: a clinicopathologic study. *Archives of Neurology*, 62, 758–765.
- Galvin, J. E., Boeve, B. F., Duda, J. E., Galasko, D. R., Kaufer, D., Leverenz, J. B., ... Lopez, O. L. (2008). Current issues in Lewy body dementia diagnosis, treatment and research. Retrieved from Lewy Body Dementia Association website: http://www.lbda.org/sites/default/files/current_issues.pdf.
- Gaugler, J. E., Ascher-Svanum, H., Roth, D. L., Fafowora, T., Siderowf, A., & Beach, T. G. (2013). Characteristics of patients misdiagnosed with Alzheimer's disease and their medication use: An analysis of the NACC-UDS database. *BMC Geriatrics*, 13. doi:10.1186/1471-2318-13-137.
- * Goldman, W. P., Price, J. L., Storandt, M., Grant, E. A., McKeel, D. W., Rubin, E. H., & Morris, J. C. (2001). Absence of cognitive impairment or decline in preclinical Alzheimer's disease. *Neurology*. doi:10.1212/WNL.56.3.361.
- * Hamilton, J. M., Salmon, D. P., Galasko, D., Delis, D. C., Hansen, L. A., Masliah, E., et al. (2004). A comparison of episodic memory deficits in neuropathologically-confirmed dementia with Lewy bodies and Alzheimer's disease. *Journal of the International Neuropsychological Society*. doi:10.1017/S1355617704105043.
- Hamilton, J. M., Salmon, D. P., Galasko, D., Raman, R., Emond, J., Hansen, L. A., et al. (2008). Visuospatial deficits predict rate of cognitive decline in autopsy-verified dementia with Lewy bodies. *Neuropsychology*. doi:10.1037/a0012949.
- Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2004.02.001.
- Hirni, D. I., Kivisaari, S. L., Monsch, A. U., & Taylor, K. I. (2016). Distinct neuroanatomical bases of episodic and semantic memory performance in Alzheimer's disease. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2013.01.013.
- Hulette, C. M., Welsh-Bohmer, K. A., Murray, M. G., Saunders, A. M., Mash, D. C., & McIntyre, L. M. (1998). Neuropathological and neuropsychological changes in "normal" aging: evidence for preclinical Alzheimer disease in cognitively normal individuals. *Journal of Neuropathology and Experimental Neurology*. doi:10.1097/00005072-199812000-00009.
- Hutchinson, A. D., & Mathais, J. L. (2007). Neuropsychological deficits in frontotemporal dementia and Alzheimer's disease: a meta-analytic review. *Journal of Neurology, Neurosurgery, and Psychiatry*. doi:10.1136/jnnp.2006.100669.
- Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., et al. (2012). National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & Dementia*, 8, 1–13.
- Jacobs, D. M., Sano, M., Dooneief, G., Marder, K., Bell, K. L., & Stern, Y. (1995). Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology*. doi:10.1212/WNL.45.5.957.
- John, S. E., Gurnani, A. S., Bussell, C. A., Saurman, J. L., Griffin, J. W., & Gavett, B. E. (in press). The effectiveness and unique contribution of neuropsychological tests and the δ latent dementia phenotype in the differential diagnosis of dementia in the uniform data set. *Neuropsychology*.
- Johns, E. K., Phillips, N. A., Belleville, S., Goupil, D., Babins, L., Kelner, L., et al. (2009). Executive functions in frontotemporal dementia and Lewy body dementia. *Neuropsychology*. doi:10.1037/a0016792.
- Johnson, D. K., & Galvin, J. E. (2011). Longitudinal changes in cognition in Parkinson's disease with and without dementia. *Dementia and Geriatric Cognitive Disorders*. doi:10.1159/000323570.
- Johnson, D. K., Morris, J. C., & Galvin, J. E. (2005). Verbal and visuospatial deficits in dementia with Lewy bodies. *Neurology*. doi:10.1212/01.wnl.0000180964.60708.c2.

- * Johnson, D. K., Storandt, M., Morris, J. C., Langford, Z. D., & Galvin, J. E. (2008). Cognitive profiles in dementia: Alzheimer disease vs healthy brain aging. *Neurology*. doi:10.1212/01.wnl.0000335972.35970.70.
- Johnson, D. K., Storandt, M., Morris, J. C., & Galvin, J. E. (2009). Longitudinal study of the transition from healthy aging to Alzheimer's disease. *Archives of Neurology*. doi:10.1001/archneur.1009.158.
- Josephs, K. A., Tsuboi, Y., Cookson, N., Watt, H., & Dickson, D. W. (2004). Apolipoprotein E ϵ 4 is a determinant for Alzheimer-type pathologic features in tauopathies, synucleinopathies, and frontotemporal degeneration. *Archives of Neurology*. doi:10.1001/archneur.61.10.1579.
- Kanne, S. M., Balota, D. A., Storandt, M., McKeel, D. W., & Morris, J. C. (1998). Relating anatomy to function in Alzheimer's disease: neuropsychological profiles predict regional neuropathology 5 years later. *Neurology*. doi:10.1212/WNL.50.4.979.
- Kazee, A. M., Eskin, T. A., Lapham, L. W., Gabriel, K. R., McDaniel, K. D., & Hamill, R. W. (1993). Clinicopathologic correlates in Alzheimer disease: assessment of clinical and pathologic diagnostic criteria. *Alzheimer Disease & Associated Disorders*, 7, 152–164.
- Khachaturian, Z. S. (1985). Diagnosis of Alzheimer's disease. *Archives of Neurology*, 42, 1097–1105. doi:10.1001/archneur.1985.04060100083029.
- Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*. doi:10.1002/jrsm.35.
- Kraybill, M. L., Larson, E. B., Tsuang, D. W., Teri, L., McCormick, W. C., Bowen, J. D., et al. (2005). Cognitive differences in dementia patients with autopsy-verified Alzheimer's disease, Lewy body pathology, or both. *Neurology*, 64, 2069–2073.
- Lamar, M., Price, C. C., Davis, K. L., Kaplan, E., & Libon, D. J. (2002). Capacity to maintain mental set in dementia. *Neuropsychologia*. doi:10.1016/S0028-3932(01)00125-7.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.
- Lopez, O. L., Wisniewski, S. R., Becker, J. T., Boller, F., & DeKosky, S. T. (1999). Psychiatric medication and abnormal behavior as predictors of progression in probable Alzheimer disease. *Archives of Neurology*, 56, 1266–1272.
- Mathais, J. L., & Burke, J. (2009). Cognitive functioning in Alzheimer's and vascular dementia. *Neuropsychology*. doi:10.1037/a0015384.
- McKeith, I. G., Burn, D., O'Brien, J., Perry, R., & Perry, E. (2002). Dementia with Lewy bodies. In K. L. Davis, D. Charney, J. T. Koyle, & C. Nemeroff (Eds.), *Neuropsychopharmacology: The fifth generation of progress* (pp. 1301–1315). Brentwood, TN: American college of neuropsychopharmacology.
- McKeith, I. G., Dickson, D. W., Lowe, J. J., Emre, M. M., O'Brien, J. T., Feldman, H. H., et al. (2005). Diagnosis and management of dementia with Lewy bodies: third report of the DLB consortium. *Neurology*. doi:10.1212/01.wnl.0000187889.17253.b1.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr., C. R., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. doi:10.1016/j.jalz.2011.03.005.
- Meehl, P. E. (1995). Bootstraps taxometrics: solving the classification problem in psychopathology. *American Psychologist*. doi:10.1037/0003-066X.50.4.266.
- * Minoshima, S., Foster, N. L., Sima, A. A., Frey, K. A., Albin, R. L., & Kuhl, D. E. (2001). Alzheimer's disease versus dementia with Lewy bodies: cerebral metabolic distinction with autopsy confirmation. *Annals of Neurology*. doi:10.1002/ana.1133.
- 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*. doi:10.1016/j.jpsychires.2008.04.014.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6(7), e1000097. doi:10.1371/journal.pmed1000097.
- Mollenhauer, B., Förstl, H., Deuschl, G., Storch, A., Oertel, W., & Trenkwalder, C. (2010). Lewy body and Parkinsonian dementia: common but often misdiagnosed conditions. *Deutsches Ärzteblatt International*. doi:10.3238/arztebl.2010.0684.
- Mrak, R. E., & Griffin, W. S. T. (2007). Dementia with Lewy bodies: definition, diagnosis, and pathogenic relationship to Alzheimer's disease. *Neuropsychiatric Disease and Treatment*, 3, 619–625.
- Mungas, D., Beckett, L., Harvey, D., Tomaszewski Farias, S., Reed, B., Carmichael, O., et al. (2010). Heterogeneity of cognitive trajectories in diverse older persons. *Psychology and Aging*. doi:10.1037/a0019502.
- Newell, K. L., Hyman, B. T., Growdon, J. H., & Hedley-Whyte, E. T. (1999). Application of the National Institute on Aging (NIA)-Reagan institute criteria for the neuropathological diagnosis of Alzheimer disease. *Journal of Neuropathology & Experimental Neurology*. doi:10.1097/00005072-199911000-00004.
- Palmer, K., Bäckman, L., Winblad, B., & Fratiglioni, L. (2003). Detection of Alzheimer's disease and dementia in the preclinical phase: population based cohort study. *British Medical Journal*. doi:10.1136/bmj.326.7383.245.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
- Price, D. L., Tanzi, R. E., Borchelt, D. R., & Sisodia, S. S. (1998). Alzheimer's disease: genetic studies and transgenic models. *Annual Review of Genetics*. doi:10.1146/annurev.genet.32.1.461.
- Price, J. L., McKeel, D. W., Buckles, V. D., Roe, C. M., Xiong, C., Grundman, M., et al. (2009). Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease. *Neurobiology of Aging*. doi:10.1016/j.neurobiolaging.2009.04.002.
- Reed, B. R., Mungas, D. M., Kramer, J. H., Ellis, W., Vinters, H. V., Zarow, C., et al. (2007). Profiles of neuropsychological impairment in autopsy-defined Alzheimer's disease and cerebrovascular disease. *Brain*. doi:10.1093/brain/awl385.
- * Riley, K. P., Jicha, G. A., Davis, D., Abner, E. L., Cooper, G. E., Stiles, N., et al. (2011). Prediction of preclinical Alzheimer's disease: longitudinal rates of change of cognition. *Journal of Alzheimer's Disease*. doi:10.3233/JAD-2011-102133.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*. doi:10.1037/0033-2909.86.3.638.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*. doi:10.1146/annurev.psych.52.1.59.
- Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychological Methods*, 13, 19–30. doi:10.1037/1082-989X.13.1.19.
- Salmon, D. P. (2000). Disorders of memory in Alzheimer's disease. In L. S. Cermak (2nd Ed.), *Handbook of neuropsychology*, Vol. 2: Memory and its disorders (pp.155–195). Amsterdam: Elsevier.
- Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology*. doi:10.1146/annurev.psych.57.102904.190024.
- * Scheff, S. W., Price, D. A., Schmitt, F. A., Scheff, M. A., & Mufson, E. J. (2011). Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease*. doi:10.3233/JAD-2011-101782.
- * Schmitt, F. A., Davis, D. G., Wekstein, D. R., Smith, C. D., Ashford, J. W., & Markesbery, W. R. (2000). Preclinical. *AD revisited*:

- Neuropathology of Cognitively Normal Older Adults. Neurology.* doi:10.1212/WNL.55.3.370.
- Sloane, P. D., Zimmerman, S., Suchindran, C., Reed, P., Wang, L., Boustani, M., & Sudha, S. (2002). The public health impact of Alzheimer's disease, 2000-2050: potential implication of treatment advances. *Annual Review of Public Health.* doi:10.1146/annurev.publhealth.23.100901.140525.
- Storandt, M., Botwinick, J., Danziger, W. L., Berg, L., & Hughes, C. P. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. *Archives of Neurology.* doi:10.1001/archneur.1984.04050170043013.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry.* doi:10.1001/archpsyc.63.8.916.
- Tierney, M. C., Snow, W. G., Reid, D. W., Zorzitto, M. L., & Fisher, R. H. (1987). Psychometric differentiation of dementia: replication and extension of the findings of Storandt and coworkers. *Archives of Neurology.* doi:10.1001/archneur.1987.00520190032013.
- Tierney, M. C., Szalai, J. P., Snow, W. G., Fisher, R. H., Nores, A., Nadon, G., et al. (1996). Prediction of probable Alzheimer's disease in memory-impaired patients: A prospective longitudinal study. *Neurology.* doi:10.1212/WNL.46.3.661.
- * Tiraboschi, P., & Guerra, U. P. (2010). How to distinguish dementia with Lewy bodies from Alzheimer's disease? *The Open Nuclear Medicine Journal.* doi:10.2174/1876388X01002020058.
- Tiraboschi, P., Salmon, D. P., Hansen, L. A., Hofstetter, R. C., Thal, L. J., & Corey-Bloom, J. (2006). What best differentiates Lewy body from Alzheimer's disease in early-stage dementia? *Brain.* doi:10.1093/brain/awh725.
- Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society.* doi:10.1111/j.1532-5415.1992.tb01992.x.
- Varma, A. R., Snowden, J. S., Lloyd, J. J., Talbot, P. R., Mann, D. M. A., & Neary, D. (1999). Evaluation of the NINCDS-ADRDA criteria in the differentiation of Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery & Psychiatry.* doi:10.1136/jnnp.66.2.184.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48 Retrieved from <http://www.jstatsoft.org/v36/i03>.
- Walker, Z., McKeith, I., Rodda, J., Qassem, T., Tatsch, K., Booij, J., et al. (2012). Comparison of cognitive decline between dementia with Lewy bodies and Alzheimer's disease: a cohort study. *British Medical Journal Open.* doi:10.1136/bmjopen-2011-000380.
- Wang, B. W., Lu, E., Mackenzie, I. R. A., Assaly, M., Jacova, C., Lee, P. E., et al. (2012). Multiple pathologies are common in Alzheimer patients in clinical trials. *The Canadian Journal of Neurological Sciences.* doi:10.1017/S0317167100015316.
- Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology.* doi:10.1001/archneur.1991.00530150046016.
- * Yoshizawa, H., Vonsattel, J. P. G., & Honig, L. S. (2013). Early neuropsychological discriminants for Lewy body disease: an autopsy series. *Journal of Neurology, Neurosurgery, and Psychiatry.* doi:10.1136/jnnp-2012-304381.
- Zakzanis, K. K. (2001). 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. *Archives of Clinical Neuropsychology.* doi:10.1016/S0887-6177(00)00076-7.