Journal of Clinical and Experimental Neuropsychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713657736

Quantitative Evidence for Neuroanatomic and Neuropsychological Markers in Dementia of the Alzheimer's Type

Konstantine K. Zakzanis

To cite this Article: Konstantine K. Zakzanis , 'Quantitative Evidence for Neuroanatomic and Neuropsychological Markers in Dementia of the Alzheimer's Type', Journal of Clinical and Experimental Neuropsychology, 20:2, 259 - 269
To link to this article: DOI: 10.1076/jcen.20.2.259.1174
URL: http://dx.doi.org/10.1076/jcen.20.2.259.1174

Journal of Clinical and Experimental Neuropsychology
1998, Vol. 20, No. 2, pp. 259-269

1380-3395/98/2002-259$12.00

Quantitative Evidence for Neuroanatomic and Neuropsychological Markers in Dementia of the Alzheimer's Type*

Konstantine K. Zakzanis
York University, Toronto, Canada and Baycrest Center for Geriatric Care, Toronto, Canada

ABSTRACT

Meta-analytic methods were used to determine the sensitivity of neuropsychological, structural, and physiological measures to temporal-hippocampal system function in dementia of the Alzheimer's type. Effect sizes are reported for the California Verbal Learning Test, the Wechsler Memory Scale-Revised, structural (i.e., magnetic resonance imaging [MRI]), and functional (i.e., positron emission tomography [PET], single photon emission computed tomography [SPECT]) neuroimaging methods. Overall, effect sizes from MRI studies are larger than those obtained from SPECT and PET, respectively, but not as large as those obtained from the neuropsychological measures. On the basis of this finding, the neuropsychological and gross pathologic similarities between Alzheimer's disease, other dementing conditions, and mixed dementias, warrants the coupling of neuropsychological evaluation for its sensitivity with neuroimaging visualization for its specificity in improving diagnostic and differential accuracy.

One of the most researched and established hypotheses regarding etiology in all of neuroscience posits involvement of the temporal-hippocampal formation as a core deficit in dementia of the Alzheimer's type (DAT; see Cummings & Benson, 1992; Hyman, Van Hoesen, Damasio, & Barnes, 1984; O'Brien, 1995; Parks, Zec, & Wilson, 1993). It has been shown that bilateral damage to the hippocampal formation in humans leads to a profound amnesia (Corkin, 1968; Damasio, 1984; Scoville & Milner, 1957; Zola-Morgan, Squire, & Amarel, 1986). In keeping with the amnestic quality that is characteristic of the patient with DAT, the temporal-hippocampal formation is of interest in Alzheimer's research because it is the area of the greatest and probably the earliest pathological change. Indeed, a hippocampal type amnesia is an early prominent feature of Alzheimer's disease and studies focusing on the temporal-hippocampal formation have reported an impressive ability to differentiate patients with DAT from healthy normal controls (Boller & Duyckaerts, 1997; Bondi, Salmon, & Kaszniak, 1996; Delis, 1989; Hyman, Arriagada, Van Hoesen & Damasio, 1993; Moss & Albert, 1988; Zakzanis, 1998).

These observations have spurred the application of neuroimaging like magnetic resonance imaging (MRI), positron emission tomography (PET), and single photon emission tomography (SPECT) to the problem of indexing the integrity of the temporal-hippocampal formation in DAT (Albert & Lafleche, 1991; Benson, 1984; Sawle, 1995), and in aiding with differential diagnoses (see Cummings & Benson, 1992). Moreover, although caution must be exercised in extrapolating from behavioral tests to neuroanatomy (Freedman, 1994), the development and use

* Appreciation is expressed to Dr. Mark W. Bondi and two anonymous reviewers for their helpful comments during manuscript preparation.
Address correspondence to: Konstantine K. Zakzanis, York University, Department of Psychology, 4700 Keele Street, Toronto, Canada, M3J 1P3. E-mail: zakzanis@yorku.ca.
Accepted for publication: February 26, 1998.

of neuropsychological tasks such as the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987) and the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987) have enabled clinicians to further index the functional consequence of a deficient temporal-hippocampal system. By deficiency, it is meant that neuropsychological and neuroanatomical investigations of temporal-hippocampal function and structure have studied that which is less than what is considered ''normal'' in comparison to healthy controls who do not present clinically with the neurobehavioral features of DAT. Thus, although the accumulation of research findings clearly demonstrates temporal-hippocampal deficit in many patients, the sensitivity of neuroimaging and neuropsychological measures to DAT have not been reviewed, integrated, and contrasted quantitatively. In view of the increasing accumulation of empirical evidence, a quantitative evaluation using meta-analytic principles was undertaken. Two questions were formulated to guide the review as follows:

(1) Does neuropsychological testing and neuroimaging visualization provide reliable evidence of temporal-hippocampal impairment in patients with DAT, and what is the average magnitude of difference between patients and healthy controls on specific indices?

(2) Are the corresponding effect sizes large enough to be considered markers for DAT?

METHODS

Meta-analysis

Standard meta-analytic techniques were employed to synthesize the structural and functional temporal-hippocampal literature in DAT (see Cooper & Hedges, 1994; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Hunter, Schmidt, & Jackson, 1982; Rosenthal, 1991, 1995; Wolf, 1986). In addition to solving problems with traditional narrative reviews (see Wolf), meta-analysis provides tools for the analysis of magnitude. Magnitude can be indexed with the effect size estimate *d* that is meant to reflect the degree to which the dependent variable is present in a sample group (e.g., a neuropsychological or neuroanatomic deficit) or the degree to which the null hypothesis is false (Cohen, 1988). In mathematical terms, *d* is the difference between patient and control means calibrated in pooled standard deviation units. Eligible research studies comprising a common dependent variable and statistics that can be transformed into effect sizes are viewed as a population to be systematically sampled and surveyed. Individual study results (typically means and standard deviations from each group) and moderator variables (e.g., education, duration of disease, gender composition, age) are then abstracted, quantified and coded, and assembled into a database that is statistically analyzed (Lipsey & Wilson, 1993). The main statistic presented in a meta-analysis is the mean effect size, which is meant to reflect the average individual effect size across the sample of studies included in the synthesis. Moderator variables are then correlated to the effect size in order to tease out relationships of subject characteristics that may influence the magnitude of the size of effect between the groups being compared. Moreover, the effect size can then be transformed into a non-overlap percentage (U) using Cohen's idealized distributions. The U statistic represents the degree of nonoverlap associated with *d* and the distribution of scores between groups (Cohen). The U statistic can further be converted to represent the degree of overlap by subtracting the nonoverlap from 100 (see Heinrichs & Zakzanis, in press). Where appropriate, this hypothetical overlap statistic (OL%; overlap percentage) will be mentioned to aid interpretation of the data. Accordingly, the OL% statistic used here represents the degree of overlap between patients with DAT and normal control participants in the distributions of cognitive, structural, and physiological scores and measures.

Literature Search

To locate potential studies, a computer-based search using the *PSYCHINFO* and *MEDLINE* databases was conducted. Independent key word searches included Alzheimer's with Positron Emission Tomography, PET, Magnetic Resonance Imaging, MRI, Single Photon Emission Tomography, SPECT, neuroimaging, California Verbal Learning Test, CVLT, Wechsler Memory Scale-Revised, WMS, WMS-R, memory, temporal lobe, hippocampus, and temporal-hippocampal. Articles were gathered from 1984 to 1997, and included only studies written in English. Articles were obtained at two large Canadian University libraries, a geriatric hospital library, and through interlibrary loan. Unfortunately, given the small and focused set of key words in any particular abstract or title that a computer search-engine such as *PSYCHINFO* and

MEDLINE uses to search for appropriate ''hits,'' some relevant studies that may have met criteria for inclusion into the meta-analyses will undoubtedly have been missed. Notwithstanding, a number of meta-analyses have been published (e.g., Raz & Raz, 1990) that have employed a computer-based search method. Indeed, whether a review is narrative or quantitative, it is difficult to be completely convinced of comprehensiveness given limited computer-search engines and the ''file-drawer'' phenomenon (see Rosenthal, 1991).

Criteria for Inclusion

Articles were included in the meta-analyses if they met or contained the following criteria: (1) publication between 1984 and 1997, (2) statistics that could be converted to effect sizes (e.g., means, standard deviations, *F*, *t*, *X*, *p*; see Johnson, 1989; Wolf, 1986), (3) patients that met NINCDS-ADRDA diagnostic criteria for ''probable DAT (McKhann et al., 1984) and, (4) a matched normal healthy control group was incorporated into the research design. The year 1984 was chosen as a year of publication criterion because it corresponded to the introduction and use of more systematic and reliable diagnostic criteria for DAT as proposed by the NINCDS-ADRDA group (McKhann et al., 1984). The year 1997 was chosen as an upper year limit to ensure full coverage of the literature by the computer journal data base which typically lags in publication date. If these criteria were met, the article was then scrutinized for two further criteria. First, the neuropsychological tests must have been administered by an examiner trained in standardized testing procedures and supervised by a psychologist. Second, the reading of MRI, SPECT, and PET scans must have been made by a researcher who was blind to the subject's condition. This was important to track in order to insure that the quality of the neuropsychological and neuroimaging evaluation in each study was held relatively constant and did not influence the findings in any systematic or a priori manner. If the research article met the above criteria, its content variable(s) was (were) included in one or more of the meta-analyses. In the case of separately published studies that used the same subject samples, the decision was adopted to treat these studies as a single study with multiple independent variables (see Hedges & Olkin, 1985). The *d* statistic was calculated for each comparison as the difference between Alzheimer's and control-group means normalized by the pooled standard deviation. Whenever means and standard deviations were reported, these were used to derive effects. When inferential statistics were reported without central tendency and dispersion data, the effects were calculated from these statistics based on formulas provided by Wolf (1986) and Johnson (1989).

Recorded Variables

Recorded variables from each study included the full study reference, moderator variables (e.g., mean age, Mini-Mental State Examination total score (MMSE; Folstein, Folstein, & McHugh, 1975), Mattis Dementia Rating Scale total score (DRS; Mattis, 1988), Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) full scale IQ estimate, education, onset age, duration of illness, and any severity index; also the type of neuroimaging instrumentation along with the procedural outline—that is, length of cut, number of cuts, whether measurement of structure was blind or not blind, T1 or T2 weighted, and the statistics used to calculate the effect size—that is, means, standard deviations, *F*, *t*, *X*, or *p*.

For the neuropsychological tasks, CVLT scores gathered included total items recalled on list A trials 1, 5, and 1-5, list B, short-delay free and cued recall, long-delay free and cued recall, and the recognition-discriminability index (see, Delis et al., 1987). From the WMS-R, scores gathered included logical memory immediate and delayed recall, and visual reproduction immediate and delayed recall. These particular measures were chosen to represent the neuropsychological memory tasks because they are used extensively in Alzheimer's research, making their inclusion in the present review obligatory. Moreover, these tasks (e.g., CVLT) provide qualitative measures of memory functioning (Delis et al., 1987; Lezak, 1995) which most other tasks of memory function do not provide (e.g., Rey Auditory Verbal Learning Test; Lezak). Furthermore, it has been demonstrated that the CVLT and the WMS-R share a high degree of convergence in long-delay free recall with an obtained correlation of .93 (Delis, Cullum, Butters, & Cairns, 1988).

For the structural effect sizes, MRI effect sizes were calculated for both temporal and hippocampal volume when presented separately, and when measured as a single temporal-hippocampal structure. Both left and right volume were calculated separately and also combined to yield independent and total volumes. The major components of the temporal lobe as defined in primary studies included the hippocampus, parahippocampal gyrus, the perirhinal cortex, the amygdala, the limbic association cortices; and subcortical structures such as the cholinergic basal forebrain; the mammillary bodies; and the anterior, dorsomedial, and midline thalamic nuclei. Based on the available behavioral

evidence in humans and animals, and experimental anatomical studies, these forebrain structures have all been implicated as playing a role in the neural substrate of memory (Hyman et al., 1993). Moreover, because of differences in test measures and indices, the PET and SPECT findings were not combined and are instead presented as a range. PET and SPECT techniques that were gathered included activation and at-rest studies, blood flow studies, and glucose metabolism studies.

RESULTS AND DISCUSSION

Twenty-seven published studies met inclusion criteria for meta-analytic synthesis (see Table 1). Across studies, 619 patients with DAT and 659 normal control subjects were assessed by either neuropsychological, MRI, PET, or SPECT methods. Table 2 includes sample size statistics for each of the meta-analyses. The larger number of effect sizes derived from the smaller number of studies reflects the fact that most studies contributed more than one effect size to the mean meta-analytic effect size (e.g., a study may have reported data on left temporal lobe volume and right temporal lobe volume, thus, contributing two effect sizes to the overall mean effect size for temporal lobe volume; conversely, the same study may have contributed neuropsychological as well as neuroimaging data or both PET and MRI data; the non-independence of effect sizes from single studies was not weighted in the mean effect size; (see Rosenthal, 1995).

Table 1. Patient Characteristics from Each Study Included in the Meta-Analyses.

Study	Age *M* (years)	Males *n*	Females *n*	Education *M* (years)
Bondi et al. (1994)	71.1	14	11	14.4
Butler et al. (1995)	80.2	5	6	NR
Cullum et al. (1995)	68.0	NR	NR	13.5
Cutler (1986)	65.3	4	8	NR
Delis et al. (1992)	74.2	7	6	14.5
Duara et al. (1986)	64.0	12	9	15.0
Erkinjuntti et al. (1993)	70.0	22	17	NR
Filley et al. (1989)	71.0	33%	67%	13.0
Guze et al. (1991)	68.0	NR	NR	NR
Haut et al. (1996)	72.5	7	10	12.2
Haxby et al. (1990)	65.7	19	13	14.8
Jernigan et al. (1991)	70.0	14	11	13.8
Jobst et al. (1992)	NR	NR	NR	NR
Johnson et al. (1988)	71.7	12	25	NR
Kessiak et al. (1991)	72.3	5	3	NR
Kessler et al. (1991)	65.4	6	15	NR
Killiany et al. (1993)	72.0	2	6	NR
Kohler (1994)	75.5	5	12	NR
Kumar et al. (1991)	67.3	NR	NR	NR
Meyer et al. (1995)	69.0	7	6	NR
Pearlson et al. (1992)	73.4	47%	53%	NR
Rabins et al. (1991)	NR	NR	NR	NR
Scheltens et al. (1992)	72.8	2	19	NR
Schlosser et al. (1989)	75.1	NR	NR	8.5
Schmidt (1992)	68.2	5	22	NR
Simon et al. (1994)	72.5	7	6	11.2
Szelies et al. (1994)	60.7	10	14	NR

Note. The majority of studies did not provide data regarding age of onset and duration of illness; see text for mean values based on 5 data points. NR = not reported.

It is apparent from Table 2 that published quasi-experimental research on Alzheimer's disease that primarily investigates the integrity of the temporal-hippocampal formation is typically conducted with neuropsychological, MRI, and PET instrumentation. Literature that met inclusion criteria for SPECT is lacking, which does not necessarily imply that the method is not widely utilized in the differential diagnosis of DAT. It may simply reflect the lack of published studies that couple a quasi-experimental methodological design with SPECT instrumentation.

Table 3 includes descriptive statistics from the 27 published studies incorporating the 619 patients with DAT and the 659 healthy controls. The median, mean, standard deviation, range, and the number of studies in which the moderator variable was reported are included in the table. Patients with DAT were typically female, 70 years old, and had suffered from the disease for approximately 4 years. Mentally, patients were moderately impaired as the average MMSE score was 18.8. Their mean education level was approximately 13 years (see Table 1 for primary study demographic data for the patients with DAT).

Moderator variable analysis is an essential part of meta-analytic investigation because it assesses the contribution of study differences in sample attributes, design, and instrument features to effect size, heterogeneity, and replication (see Heinrichs & Zakzanis, in press). From Tables 1 and 3 it is evident that variability is present in which the rate of clinical and demographic patient characteristics are reported. In terms of moderator variables, the published literature on neurocognition and neuroimaging in Alzheimer's disease is surprisingly limited and often inadequate. Basic subject attributes like age and gender composition were not reported in all studies. Clinical variables of special relevance to DAT, like the approximate duration of disease, mental status, and severity of disease were so underreported that a limited correlational study of potential moderators of effect was conducted, with many correlations missing statistical significance because of inadequate power. This lack of reported data limits the validity of the present synthesis since it can not be comfortably concluded that the subject samples from the primary studies are drawn from a single underlying population, although all patients did indeed meet NINCDS-ADRDA diagnostic criteria for probable DAT. Unfortunately, the analyses based on a demographic data set with so many missing variables compromises the basic assumption that subjects from the primary studies share common demographic and clinical attributes. Thus, the present synthesis should be considered a ''preliminary'' meta-analytic investigation of neuropsychological and neuroanatomic markers in DAT until it can be concluded firmly that the obtained effect sizes are evidence of deficient temporal-hippocampal function and structure across various samples of patients with DAT. Such evidence will depend on the individual reporting practices of investigators and, presumably, the expectations of journal editors and reviewers.

From the neuropsychological effects, significant correlations were found for education (r =

Table 2. Sample Size Statistics.

Method	*n*	*nd*	DAT *n*	Con. *n*
Neuropsychological Assessment (CVLT & WMS-R)	11	43	184	244
MRI	9	34	170	204
PET	8	12	191	177
SPECT	3	6	74	34
Total across methods	31	95	619	659

Note. CVLT = California Verbal Learning Test; WMS-R = Wechlser Memory Scale-Revised; MRI = magnetic resonance imaging; PET = positron emission tomography; SPECT = single photon emission tomography; *n* = number of studies; *nd* = number of effect sizes; DAT *n* = number of patients with Alzheimer's disease; Con. *n* = number of normal healthy controls.

Table 3. Descriptive Statistics for Temporal-Hippocampal Studies of Alzheimer's Disease.

Variables	*Mdn*	*M*	(*SD*)	Range	*n*
Sample size					
DAT samples	20	23	(11.0)	8.0 – 51.0	27
Control samples	16	25	(15.5)	7.0 – 65.0	27
Patients' age	71.0	70.2	(4.2)	60.7 – 80.2	25
Percentage of females	56.5	54.0	(22.4)	0.0 – 90.0	21
Patients' education (yrs.)	13.6	13.1	(2.0)	8.5 – 15.0	10
Patients' MMSE	19.1	18.8	(2.9)	14.0 – 23.9	18
Patients' DRS	96.0	99.8	(14.0)	83.0 –113.8	5
Patients' WAIS-R IQ	95.8	95.8	(3.9)	93.0 – 98.5	2

Note. DAT = dementia of the Alzheimer's type; (yrs.) = years; MMSE = Mini-Mental State Exam; DRS = Mattis Dementia Rating Scale; WAIS-R IQ = Wechsler Adult Intelligence Scale-Revised Intelligence Quotient. *n* = number of papers in which the moderator was reported.

–.41, $p < .05$, $n = 10$), percentage of females ($r =$ –.59, $p < .005$, $n = 21$), and *d*. The negative relationship between education and *d* is hard to interpret with such a limited sample size, and therefore, can not be elaborated on. However, it could be speculated that higher education is related to better performance on the CVLT and the WMS-R which results in smaller effect sizes. The negative relationship between gender and *d* suggests that as more females are included in the Alzheimer's sample, the difference in neuropsychological performance on CVLT and WMS-R measures between patients with DAT and healthy controls will decrease. This may reflect a gender difference in the neurobehavioral severity of memory functioning in Alzheimer's disease where male patients are quantitatively more impaired on tasks of memory function such as the CVLT and the WMS-R. In keeping with the significant relationship between education and the neuropsychological tasks, this gender discrepancy in severity of memory impairment might reflect the higher education levels typically achieved by females.

From the SPECT effects, the mean MMSE was significantly related to the mean meta-analytic effect size ($r = .95$, $p < .05$, $n = 5$). Again, however, the relationship is difficult to interpret with such a limited sample size of SPECT studies that incorporated the MMSE. In terms of significant moderators of effect including neuroimaging variables, there were no other significant product-moment correlations between basic demographic and clinical variables and the effect sizes.

Table 4 summarizes the mean meta-analytic effect sizes. The table includes the number of effect sizes that were used in the calculation of the mean effect size, the mean effect size *d*, the standard deviation of *d*, and the 95% confidence interval. The table also includes the hypothetical OL% based on Cohen's (1988) idealized population distributions. Again, the percentage is meant to reflect the degree of overlap in sample score dispersion associated with *d*. Thus, as the percentage of overlap decreases, the discriminability of the two groups being compared increases. For example, if hippocampal atrophy is a core marker found in all patients with Alzheimer's disease, an effect size capable of completely discriminating patients from healthy controls is a defensible expectation. Such an effect size would have to be about 3.0 or greater, as the corresponding overlap (OL%) associated with the effect is less than approximately 5%. With an effect size considerably smaller (i.e., corresponding overlap of say, > 15%), it would be hard to argue in favor of hippocampal atrophy as being a core marker of Alzheimer's dementia. Although such a standard is not entirely justifiable, it can serve as a heuristic benchmark in which the present findings can be articulated.

From Table 4, it can be seen that the effect size results indicate that the CVLT and WMS-R

Table 4. Effect Size Results.

Method	*nd*	*Md*	(*SDd*)	OL%	95%CI
Neuropsychological Task					
CVLT	26	3.3	(1.4)	5.0	2.8–3.8
WMS-R	17	2.8	(1.1)	8.8	2.2–3.6
Combined	43	3.2	(1.1)	5.8	2.6–3.6
MRI	34	1.4	(.96)	31.9	1.0–1.7
PET	12	1.2	(.77)	37.8	.69–1.7
SPECT	6	1.4	(.58)	31.9	.95–2.2

Note. Effect sizes are displayed in absolute values to aid comparison. *nd* = number of effect sizes; *Md* = mean meta-analytic effect size; (SD*d*) = standard deviation of *d*; OL% = percent overlap based on Cohen's (1988) idealized population distributions; 95%CI = 95% confidence interval. CVLT = California Verbal Learning Test; WMS-R = Wechlser Memory Scale-Revised; MRI = magnetic resonance imaging; PET = positron emission tomography; SPECT = single photon emission tomography.

neuropsychological tasks were more sensitive than the neuroimaging methods to discriminating patients with DAT from normal healthy controls. Moreover, the magnitude of effect size from the neuropsychological variables met the proposed benchmark criteria for neuropsychological and neuroanatomic markers (i.e., $d > 3.0$ OL% < 5). In other words, the performance of patients with DAT on the neuropsychological variables may serve as markers for the disease, as the associated overlap in the distribution of scores between patients and controls is less than approximately 5%.

The method next most sensitive to discriminating patients from controls was MRI, followed closely by SPECT, and PET, respectively. The sensitivity of these methods in indexing temporal-hippocampal integrity does not typically enable the clinician to reliably discriminate patients with DAT from normal healthy controls when used alone. Although the neuroimaging effect sizes are large based on Cohen's (1988) idealized distributions, the corresponding overlap values range from 32% to 38%. Accordingly, the neuroimaging effects do not meet the proposed heuristic benchmark criteria for marker eligibility. In terms of sensitivity to DAT, hippocampal volume, temporal lobe volume, temporal regional cerebral metabolic rate of glucose, and temporal regional cerebral blood flow as indexed with the degree of resolution obtained with current imaging technology, all correspond to effect sizes that indicate that volumetric and metabolic rate measures of these structures can be within normal limits.

Table 5 is presented in order to illustrate a more qualitative analysis of the obtained effect sizes. The table includes the mean effect size, the corresponding OL% and the standard deviation of *d*. The variable with the most discriminating power was the CVLT long-delay free recall measure ($d = 4.47$; OL% < 2) followed by the CVLT long-delay cued recall ($d = 4.02$; OL% = 2.3), short-delay recall ($d = 3.90$; OL% = 2.8), and short-delay cued recall ($d = 3.47$; OL% = 4.5), respectively. Other variables that met benchmark criteria included the CVLT list A trials 1–5 measure, the WMS-R logical memory delayed recall and the visual reproduction delayed recall measures, the CVLT recognition-discriminability index, and finally the CVLT list A trial 5 measure (see Table 5). The sole neuroanatomic variable that approximated benchmark criteria was whole hippocampal volume as measured with MRI ($d = 2.88$; OL% = 7.7).

CONCLUSION

Consideration of the effect sizes indicate that both neuropsychological and structural and functional imaging literatures are distinguished by large effect sizes whereby most patients show

Table 5. Detailed Effect Size Analysis.

Method/variable	*Md*	(*SDd*)	OL%
Neuropsychological task			
CVLT			
list A trial 1	2.05	(———)	18.9
list A trial 5	2.94	(0.54)	8.0
list A trials 1-5	3.40	(1.08)	4.7
list B	2.66	(———)	10.4
short-delay cued recall	3.47	(0.99)	4.2
short-delay free recall	3.90	(0.43)	2.8
long-delay cued recall	4.02	(0.42)	2.3
long-delay free recall	4.47	(1.59)	< 2
recognition-discriminability	3.00	(0.92)	7.2
WMS-R			
Logical Memory			
immediate recall	2.59	(0.93)	10.7
delayed recall	3.36	(1.32)	4.7
Visual Reproduction			
immediate recall	2.34	(0.97)	14.0
delay recall	3.00	(0.92)	7.2
MRI			
whole temporal lobe volume	1.18	(0.47)	38.0
left temporal lobe volume	1.51	(0.51)	29.3
right temporal lobe volume	0.95	(0.62)	46.5
whole hippocampus volume	2.88	(3.83)	7.7
left hippocampus volume	1.11	(0.32)	41.1
right hippocampus volume	1.57	(0.97)	27.1
PET			
rCBF hippocampus	0.62	(0.66)	60.8
rCBF whole temporal lobe			
(at rest)	0.90	(0.74)	48.4
(activated)	1.64	(1.64)	26.0
rCBF left temporal lobe	1.73	(0.81)	24.0
rCBF right temporal lobe	1.67	(0.96)	25.6
SPECT			
whole temporal-hippocampal	1.31	(0.14)	34.7
left temporal-hippocampal	1.71	(1.03)	24.6
right temporal-hippocampal	1.67	(0.64)	25.6

Note. Effect sizes are displayed in absolute values to aid comparison; CVLT = California Verbal Learning Test; WMS-R = Wechsler Memory Scale—Revised; MRI = magnetic resonance imaging; PET = positron emission tomography; SPECT = single photon emission tomography; rCBF = regional cerebral blood flow; *Md* = mean effect size *d*; (*SDd*) = standard deviation of *d*; OL% = percent overlap based on Cohen's (1988) idealized distributions.

deficient temporal-hippocampal function and structure. Moreover, the mean meta-analytic effect sizes from neuropsychological studies where patients are compared to healthy controls are largest, followed by MRI, SPECT, and PET studies, respectively. Thus, the neuropsychological variables are most sensitive to indexing temporal-hippocampal dysfunction in patients with DAT. Moreover, the meta-analysis of the neuropsychological literature resulted in effect sizes for specific task variables that met the proposed heuristic benchmark criteria (i.e., $d > 3.0$ OL% < 5) for clinical neuropsychological and neuroanatomic markers.

Indeed, the nature of a patient's memory impairment can vary greatly depending on the type

and location of the neuropathology (Delis, 1989), and can therefore not serve as a sole index for diagnosis and as a differential marker for DAT. Moreover, poor performance on memory tasks might reflect the many other features of Alzheimer's disease as well. It is true that other patient populations demonstrate memory impairments such as alcoholic Korsakoff syndrome, herpes simplex encephalitis, and Huntington's disease, with varying loci of neuropathology (see Cummings & Bension, 1992). Therefore, deficient performance on neuropsychological tasks of memory function do present in many conditions besides Alzheimer's disease. Thus, although the poor performance of patients with DAT on neuropsychological tasks may serve as markers for the disease, the specificity of neuroimaging techniques to temporal-hippocampal alterations in conjunction with a complete neurological work-up is essential in the accurate diagnosis and differential of DAT from other dementing conditions with prominent features of memory dysfunction. The neuropsychological and gross pathologic similarities between Alzheimer's disease, other dementing conditions, and mixed dementias, warrants the coupling of neuropsychological evaluation for its sensitivity with neuroimaging visualization for its specificity in improving diagnostic and differential accuracy.

Finally, the proposition of heuristic benchmark criteria for neuropsychological and neuroanatomic markers ($d > 3.0$ OL% < 5) can provide a basis for identifying clinical variables that can discriminate reliably between specific brain diseases and normal cognitive function. Indeed, a compilation of specific test variable effect sizes and corresponding OL% for various neurologic and psychiatric diseases might aid in the differentiation of these disorders. That is, where $d > 3.0$ and OL% < 5, the clinician and / or researcher can be assured that the test measure can reliably discriminate approximately all patients from healthy controls. Thus, by identifying further neuropsychological and neuroanatomic markers, a systematic approach to differential diagnosis can be developed.

REFERENCES

*Studies accompanied by an asterisk were used in the meta-analysis

Albert, M. S., & Lafleche, G. (1991). Neuroimaging in Alzheimer's disease. *Psychiatric Clinics of North America, 14*, 443-459.

Benson, D. F. (1984). New developments in diagnosis—CAT, PET, and NMR. *Psychiatric Annals, 14*, 192-197.

Boller, F., & Duyckaerts, C. (1997). Alzheimer disease: Clinical and anatomic aspects. In T. E. Feinberg, & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 521-544). New York: McGraw-Hill.

*Bondi, M. W., Monsch, A. U., Galasko, D., Butters, N., Salmon, D. P., & Delis, D. C. (1994). Preclinical cognitive markers of dementia of the Alzheimer's type. *Neuropsychology, 8*, 374-384.

Bondi, M. W., Salmon, D. P., & Kaszniak, A. W. (1996). The neuropsychology of dementia. In I. Grant, & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 164-199). New York: Oxford University Press.

*Butler, R. E., Costa, D. C., Greco, A., Ell, P. J., & Katona, C. L. E. (1995). Differentiation between Alzheimer's disease and multi-infarct dementia: SPECT vs. MR imaging. *International Journal of Geriatric Psychiatry, 10*, 121-128.

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

Cooper, H., & Hedges, L. V. (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal lobe excision. *Neuropsychologia, 6*, 255-265.

*Cullum, C. M., Filley, C. M., & Kozora, E. (1995). Episodic memory function in advanced aging and early Alzheimer's disease. *Journal of the International Neuropsychological Society, 1*, 100-103.

Cummings, J. L., & Benson, D. F. (1992). *Dementia: A clinical approach*. Boston: Butterworth Heinman.

*Cutler, N. R. (1986). Cerebral metabolism as measured with positron emission tomography (PET) and [f} 2-Deoxy-D-Glucose. *Progress in Neuropsychopharmacology and Biological Psychiatry, 10*, 309-321.

Damasio, A. R. (1984). The anatomic basis of memory disorders. *Seminars in Neurology, 4*, 223-225.

Delis, D. C. (1989). Neuropsychological assessment of learning and memory. In F. Boller, & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 3, pp. 3-33). Amsterdam, The Netherlands: Elsevier Science Publishers.

Delis, D. C., Cullum, C. M., Butters, N., & Cairns, P. (1988). Wechsler Memory Scale-Revised and California Verbal Learning Test: Convergence and divergence. *The Clinical Neuropsychologist, 2*, 188-195.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test*. San Antonio, TX: The Psychological Corporation.

*Delis, D. C., Massman, P. J., Butters, N., Salmon, D. P., Shear, P. K., Demandura, T., & Filoteo, J. V. (1992). Spatial cognition in Alzheimer's disease: Subtypes of global-local impairment. *Journal of Clinical and Experimental Neuropsychology, 14*, 463-477.

*Duara, R., Grady, C., Haxby, J., Sundaram, M., Cutler, N. R., Heston, L., Moore, A., Sclageter, N., Larson, S., & Rapoport, S. I. (1986). Positron emission tomography in Alzheimer's disease. *Neurology, 36*, 879-887.

*Erkinjuntti, T., Lee, D. H., Gao, F., Steenhuis, R., Eliasziw, M., Fry, R., Merskey, H., Hachinski, V. C. (1993). Temporal lobe atrophy on magnetic resonance imaging in the diagnosis of early Alzheimer's disease. *Archives of Neurology, 50*, 305-310.

*Filley, C. M., Davis, K. A., Schmitz, S. P., Stears, J. C., Heaton, R. K., Kelly, J., Culig, K. M., & Scherzinger, A. L. (1989). Neuropsychological performance and magnetic resonance imaging in Alzheimer's disease and normal aging. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 2*, 81-90.

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

Freedman, M. (1994). Frontal and parietal lobe dysfunction in depression: Delayed alternation and tactile learning deficits. *Neuropsychologia, 32*, 1015-1025.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

*Guze, B. H., Baxter Jr., L. R., Schwartz, J. M., Szuba, M. P., Mazziotta, J. C., & Phelps, M. E. (1991). Changes in glucose metabolism in dementia of the Alzheimer type compared with depression: A preliminary report. *Psychiatry Research: Neuroimaging, 40*, 195-202.

*Haut, M. W., Weber, A. M., Demarest, D., Keefover, R. W., & Rankin, E. D. (1996). Controlling for constructional dysfunction with the visual reproduction subtest of the Wechsler Memory Scale-Revised in Alzheimer's disease. *The Clinical Neuropsychologist, 10*, 309-312.

*Haxby, J. V., Grady, C. L., Koss, E., Horwitz, B., Heston, L., Schapiro, M., Friedland, R., & Rapoport, S. I. (1990). Longitudinal study of cerebral metabolic asymmetries and associated neuropsychological patterns in early dementia of the Alzheimer's type. *Archives of Neurology, 47*, 753-760.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Heinrichs, R. W., & Zakzanis, K. K. (in press). Neurocognitive deficit in schizophrenia: A quantitative review. *Neuropsychology*.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Hyman, B. T., Arriagada, P. V., Van Hoesen, G. W., & A. R. Damasio (1993). Memory impairment in Alzheimer's disease: An anatomical perspective. In R. W. Parks, R. F. Zec, & R. S. Wilson (Eds.), *Neuropsychology of Alzheimer's disease and other dementias* (pp. 138-150). New York: Oxford University Press.

Hyman, B. T., Van Hoesen, G. W., Damasio, A. R., & Barnes, C. L. (1984). Alzheimer's disease: Cell specific pathology isolates the hippocampal formation. *Science, 225*, 1168-1170.

*Jernigan, T. L., Salmon, D. P., Butters, N., & Hesselink, J. R. (1991). Cerebral structure on MRI, part II: Specific changes in Alzheimer's and Huntington's diseases. *Biological Psychiatry, 29*, 68-81.

*Jobst, K. A., Smith, A. D., Barker, C. S., Wear, A., King, E. M., Smith, A., Anslow, P. A., Molyneux, A. J., Shepstone, B. J., Soper, N., Holmes, K. A., Robinson, J. R., Hope, R. A., Oppenheimer, C., Brockbank, K., & McDonald, B. (1992). Association of atrophy of the medial temporal lobe with reduced blood flow in the posterior parietotemporal cortex in patients with a clinical and pathological diagnosis of Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*, 190-194.

Johnson, B.T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Lawerence Erlbaum.

*Johnson, K. A., Holman, B. L., Mueller, S. P., Rosen, T. J., English, R., Nagel, S., & Growdon, J. H. (1988). Single photon emission computed tomography in Alzheimer's disease: Abnormal Iofetamine I^{123} uptake reflects dementia severity. *Archives of Neurology, 45*, 392-396.

*Kessiak, J. P., Nalcioglu, O., & Cotman, C. W. (1991). Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology, 41*, 51-54.

*Kessler, J., Herholz, K., Ground, M., & Heiss, W. D. (1991). Impaired metabolic activation in Alzheimer's disease: A PET study during continuous visual recognition. *Neuropsychologia, 29*, 229-243.

*Killiany, R. J., Moss, M. B., Albert, M. S., Sandor, T., Tieman, J., & Jolesz, F. (1993). Temporal lobe

regions on magnetic resonance imaging identify patients with early Alzheimer's disease. *Archives of Neurology, 50*, 949-954.

*Kohler, S. (1994). Quantitative characterization of verbal learning deficits in patients with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 16*, 749-753.

*Kumar, A., Schapiro, M. B., Grady, C., Haxby, J. V., Wagner, E., Salerno, J. A., Friedland, R. P., & Rapoport, S. I. (1991). High-resolution PET studies in Alzheimer's disease. *Neuropsychopharmacology, 4*, 35-46.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment**.** *American Psychologist, 12*, 1181-1209.

Mattis, S. (1988). *Dementia Rating Scale.* Odessa, FL: Psychological Assessment Resources.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the Department of Health and Human Service Task Force on Alzheimer's Disease. *Neurology, 34*, 39-44.

*Meyer, M., Koeppe, R. A., Frey, K. A., Foster, N. L., & Kuhl, D. E. (1995). Positron emission tomography measures of Benzodiazepine binding in Alzheimer's disease. *Archives of Neurology, 52*, 314-317.

Moss, M. B., & Albert, M. S. (1988). Alzheimer's disease and other dementing disorders. In M. S. Albert, & M. B. Moss (Eds.), *Geriatric neuropsychology* (pp. 145-178). New York: Guilford Press.

O'Brien, J. T. (1995). Is hippocampal atrophy on magnetic resonance imaging a marker for Alzheimer's disease. *International Journal of Geriatric Psychiatry, 10*, 431-435.

Parks, R. W., Zec, R. F., & Wilson, R. S. (1993). *Neuropsychology of Alzheimer's disease and other dementias*. New York: Oxford University Press.

*Pearlson, G. D., Harris, G. J., Powers, R. E., Barta, P. E., Camargo, E. E., Chase, G. A., Noga, T., & Tune, L. E. (1992). Quantitative changes in mesial temporal volume, regional cerebral blood flow, and cognition in Alzheimer's disease. *Archives of General Psychiatry, 49*, 402-408.

*Rabins, P. V., Pearlson, G. D., Aylward, E., Kumar, & Dowell, K. (1991). Cortical magnetic resonance imaging changes in elderly inpatients with major depression. *American Journal of Psychiatry, 148*, 617-620.

Raz, S., & Raz, N. (1990). Structural brain abnormalities in the major psychoses: A quantitative review of the evidence from computerized imaging. *Psychological Bulletin, 108*, 93-108.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183-192.

Sawle, G. V. (1995). Imaging the head: Functional imaging. *Journal of Neurology, Neurosurgery, and Psychiatry, 58*, 132-144.

*Scheltens, P. H., Leys, D., Barkhof, F., Huglo, D., Weinstein, H. C., Vermersch, P., Kuiper, M., Steinling, M., Wolters, E. C. H., & Valk, J. (1992). Atrophy of medial temporal lobes on MRI in ''probable'' Alzheimer's disease and normal aging: Diagnostic value and neuropsychological correlates. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*, 967-972.

*Schlosser, D., & Ivison, D. (1989). Assessing memory deterioration with the Wechsler Memory Scale, the National Adult Reading Test, and the Schonell Graded Word Reading Test. *Neuropsychology*, 11, 785-792.

*Schmidt, R. (1992). Comparison of magnetic resonance imaging in Alzheimer's disease, vascular dementia and normal aging. *European Neurology, 32*, 164-169.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11-21.

*Simon, E, Leach, L., Winocour, G., & Moscovitch, M. (1994). Intact primary memory in mild to moderate Alzheimer's disease: Indices from the California Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology, 16*, 414-422.

*Szelies, B., Mielke, R., Herholz, K., Heiss, W. D. (1994). Quantitative topographical EEG compared to FDG PET for classification of vascular and degenerative dementia. *Electroencephalography and Clinical Neurophysiology, 91*, 131-139.

Wechsler, D. (1981). *WAIS-R [manual]*. New York: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale manual*. San Antonio, TX: The Psychological Corporation.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.

Zakzanis, K. K. (1998). Indexing temporal-hippocampal integrity in dementia of the Alzheimer's type [abstract]. *Archives of Clinical Neuropsychology, 13*, 51.

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human amnesia and the medial temporal lobe region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience, 6*, 2950-2967.