DOI: 10.1017/S1355617703950053

The neuropsychological profiles of mild Alzheimer's disease and questionable dementia as compared to age-related cognitive decline

ELISE CACCAPPOLO-VAN VLIET,[1] JENNIFER MANLY,[1,2] MING-XIN TANG,[1,4,5] KAREN MARDER,[1,2,4] KAREN BELL,[1,2,4] AND YAAKOV STERN[1–4]

[1]Gertrude H. Sergievsky Center, Columbia University College of Physicians and Surgeons, New York, New York
[2]Department of Neurology, Columbia University College of Physicians and Surgeons, New York, New York
[3]Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York
[4]Cognitive Neuroscience Division, Taub Institute for Research in Alzheimer's Disease and the Aging Brain, Columbia University College of Physicians and Surgeons, New York, New York
[5]Division of Biostatistics, Columbia University College of Physicians and Surgeons, New York, New York

(RECEIVED September 25, 2001; REVISED August 2, 2002; ACCEPTED August 5, 2002)

Abstract

Test scores from a comprehensive neuropsychological battery administered to 1602 subjects consisting of 1347 subjects with probable Alzheimer's disease (AD), 100 subjects with questionable dementia (QD) and 155 non-demented elderly control subjects were cross-sectionally analyzed. Subjects with probable AD were categorized as *mild* ($n = 244$), *moderate* ($n = 480$), *severe* ($n = 376$), and *very severe* ($n = 247$) according to modified mini mental status exam (mMMSE) scores. Mean scores on individual neuropsychological tests are provided for each group of subjects. Stratified random sampling was performed to select a sample of mild AD subjects who were matched in age and education to non-demented elderly controls, and analyses focused on the performance of QD subjects and mild AD subjects, whose scores were compared to those of the elderly control subjects. Selected scores were organized by cognitive domain and logistic regressions were used to determine the domains and individual tests within each that were most predictive of group status. Results suggested a profile of scores associated with QD and mild AD including impaired recall of verbal information for both groups. Areas of lower functioning in QD subjects as compared to elderly controls included category fluency and visuospatial ability.
(*JINS*, 2003, *9*, 720–732.)

Keywords: Alzheimer's disease, Questionable dementia, Age-related cognitive decline, Neuropsychological testing

INTRODUCTION

While the identification of risk factors associated with Alzheimer's disease (AD), such as age, education, occupational attainment, and genetic factors, has advanced our understanding of the disease, of equal importance is the ability to reliably and accurately diagnose probable AD, particularly in its earliest stages. According to the diagnostic criteria for AD, published in 1984 by a joint consensus conference between the National Institute of Neurological and Communicative Disorders (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA; McKhann et al., 1984), a research diagnosis of AD requires dementia confirmed by neuropsychological testing. Because the diagnosis of AD remains one of exclusion, the neuropsychological evaluation represents the "cornerstone" of an AD diagnosis. Furthermore, any studies examining the risk factors of AD rely on its accurate diagnosis.

Research criteria for probable AD include: onset between the ages of 40 and 90, most often after age 65; a dementia established by clinical evaluation and documented by the Mini-Mental State Examination, Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests; deficits in at least two domains of cognition; progressive decline of memory and other cognitive functions; no disturbance of consciousness; and the absence of any other disorders that could account for

Reprint requests to: Yaakov Stern, Gertrude H. Sergievsky Center, 630 West 168th Street, New York, NY, 10032. E-mail: ys11@columbia.edu

the progressive cognitive decline (McKhann et al., 1984). The clinical diagnosis of AD can be achieved with a high level of certainty by determining the presence of dementia, identifying patterns of relatively impaired and intact cognitive abilities, and eliminating other potential causes (Galasko et al., 1994). Thus the necessity of neuropsychological evaluation in the diagnostic accuracy of AD is clear and the pattern of cognitive impairment associated with AD remains a significant topic that should not be overshadowed by the increased focus on the identification of biological risk factors.

The ability to clinically detect AD in its earliest stages is particularly valuable, especially given the recent advances in pharmacological therapies designed to slow the progression of the disease. While progressive cognitive decline is the hallmark of AD, diagnosing AD in its initial stages is confounded by the gradual cognitive slowing associated with normal aging; in general, both cross-sectional and longitudinal studies report an inverse correlation between age and performance on various neuropsychological tests (Lindenberger et al., 1993; Scherr et al., 1988; Wallace & Collins, 1991). Age-related cognitive decline is usually characterized by mild impairments in a number of cognitive domains (Petersen et al., 2001), which are typically manifested as changes in episodic memory, or a reduced ability to learn new information (Craik et al., 1987; Small et al., 1999) resulting from decreased speed of central processing required for the encoding and retrieval of information (Rabinowitz & Craik, 1986). Working memory is affected (Hultsch et al., 1992; Rabinowitz & Craik, 1986), while retention usually remains preserved (Petersen et al., 1999; Small et al., 1999) as does implicit memory (Ritchie et al., 1997). Other changes associated with normal aging include mild deficits in language functioning (naming, verbal fluency), visuospatial abilities, perceptual speed and executive functioning (Rubin et al., 1998; Salthouse, 1989; Schaie, 1989).

Recently, interest has focused on a group of subjects who are not demented but demonstrate memory impairment beyond that expected for age and education that cannot be accounted for by any recognized medical or psychiatric diagnosis. Functional status may be slightly impaired or not affected at all. This condition has been variously referred to in the literature as preclinical AD, incipient dementia, isolated memory impairment, mild cognitive impairment, and questionable dementia (QD) (Collie et al., 1999; Elias et al., 2000; Flicker et al., 1991; Tierney et al., 1996). Attempts to define the clinical and cognitive profile of patients with this condition have resulted in an imprecise set of criteria including, among other items, memory complaints documented by the patient or a collateral source, a Mini Mental State Examination (MMSE) score of 24 or higher, and a mildly abnormal score on a clinical scale for the staging of dementia (Celsis, 2000). The relatively few number of studies examining the neuropsychological performance of QD report impaired memory, particularly on delayed recall tasks (Morris et al., 1991; Tierney et al., 1996) and, more recently, on paired associate learning tasks (Fowler et al., 2002). These deficits resemble those reported in normal aging studies, demonstrating the need for more in-depth analyses of the cognitive impairment associated with QD.

A number of studies have found that elderly subjects with memory deficits are at an increased risk for developing dementia (Howieson et al., 1997; Rubin et al., 1998; Tierney et al., 1996) and QD has been described as a transitional period of cognitive decline that occurs between normal aging and AD (Petersen et al., 2001). Results of autopsies performed on patients, who at time of death were characterized as QD, showed histopathological signs associated with AD (Storandt & Hill, 1989). In fact, some investigators question the distinction between QD and AD, reasoning that QD is neuropathologically identical to AD (Morris et al., 1991), while others emphasize that not all QD patients necessarily progress to AD (Daly et al., 2000; Grundman et al., 1996; Petersen et al., 1999). Celsis (2000) estimates the conversion rate from QD to AD as approximately 12% per year and points out that this rate is 10 times higher than the incidence of dementia in the general population. Consequently, the classification of QD is paramount to the ability to differentiate between normal age-related cognitive changes and those that are characteristic of QD and AD (Doody et al., 2001; Petersen et al., 2001).

A comparison of the neuropsychological performance characteristic of normal aging, QD and mild AD can assist in diagnosis, whether aiming to differentiate QD and mild AD from normal aging or from dementias of different etiologies. Earlier studies have assessed the sensitivity and specificity for differentiating AD subjects from normal elderly subjects and those "at risk" for AD within a community-based sample (Cahn et al., 1995) and compared the test scores of AD patients to those of other dementia groups (e.g., Connor et al., 1998; Gainotti et al., 1998). To date, however, no study has examined the scores from an entire diagnostic battery across a clinically based sample of non-demented elderly subjects, QD subjects, and AD patients of varying severity. The purpose of the present study was to cross-sectionally analyze a large body of neuropsychological scores obtained from archival data on these groups. One benefit of such a study lies in the diagnostic utility of providing raw neuropsychological scores for each subject group. Given the large size of the sample of AD patients alone, these scores will hopefully provide other clinicians with a framework to reference when examining older patients for whom memory impairment is a presenting problem. Of particular interest to this study are the differences in performance demonstrated by QD patients and mild AD patients as compared to non-demented elderly controls across cognitive domains as well as on individual neuropsychological tests. Analyses of these scores allow us to differentiate the cognitive performance of elderly individuals with possible age-related deficits, currently considered within the context of normal aging, from that of questionable dementia or early manifestations of AD.

METHODS

Participant Selection and Classification

Subjects were either selected from a continuous series of patients referred to our Memory Disorders Center or were recruited for a study of controls (total $N = 2762$). Figure 1 illustrates the subject selection process. All subjects were English-speaking. Those subjects for whom English was a second language were excluded ($n = 243$). Subjects were diagnosed according to strict ADRC criteria based on full clinical evaluation and extensive neuropsychological testing. All subjects were administered a modified version of the Mini-Mental State Examination (Folstein et al., 1975), the procedures for which have been reported (Mayeux et al., 1981). In addition, each subject was evaluated by a physician on measures of functional capacity including the Blessed Dementia Rating Scale (BADL; Blessed et al., 1968) and the Schwab and England Activities of Daily Living Scale (Schwab & England, 1969). These interviews were completed with the subject, an informant, or both. A physician conducted a standardized physical and neurological examination. All ancillary information, including standard laboratory measures used in dementia diagnosis (such as

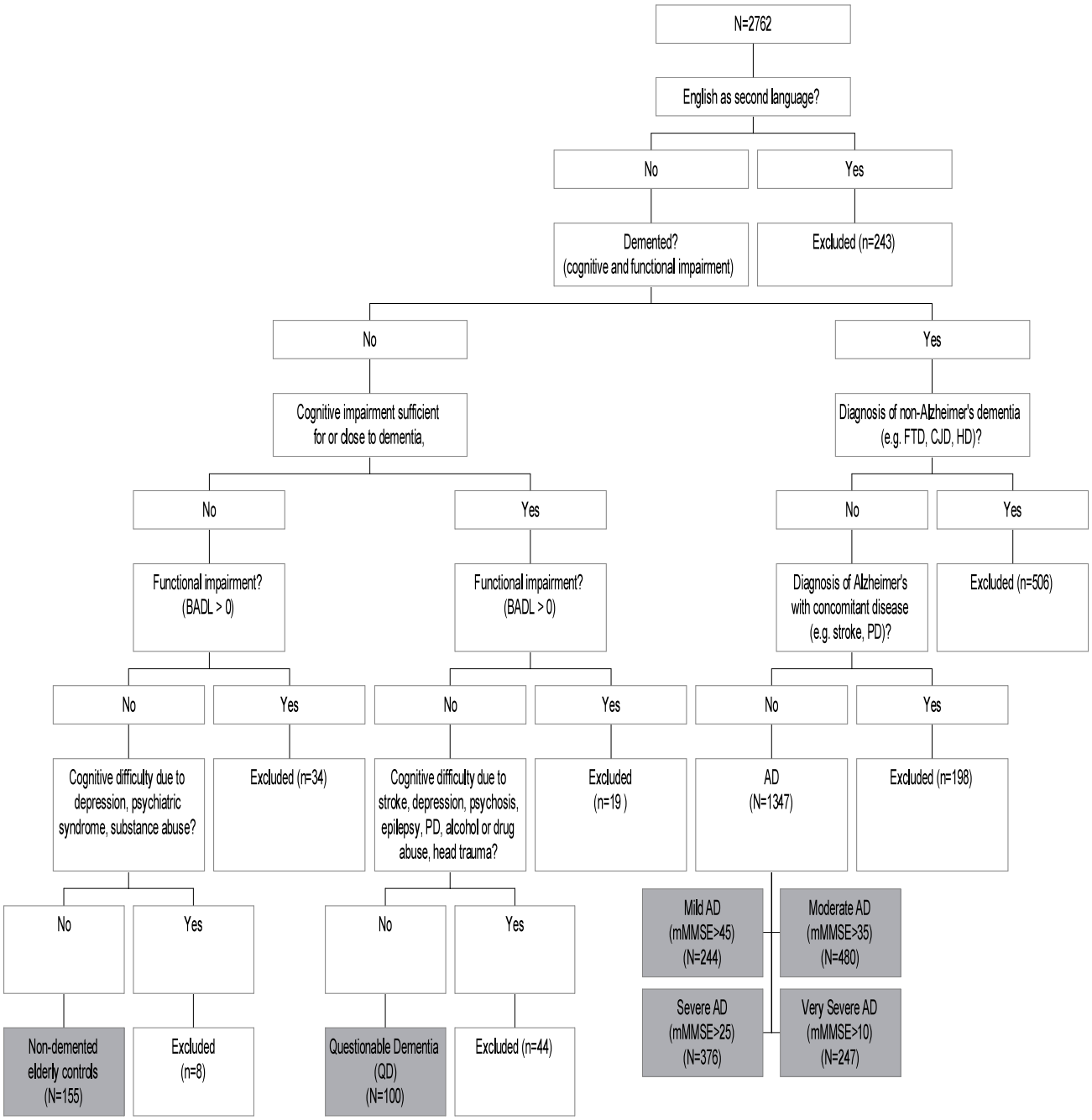

Fig. 1. Subject selection process.

chemistry, CBC, VDRL, TFTs and B_{12} levels), CT scans, MR images and functional imagings if available, were included in the evaluation. All subjects received a comprehensive neuropsychological evaluation as described below. Information from these evaluations was presented at a diagnostic conference of physicians and neuropsychologists, and a consensus diagnosis was made. The diagnosis of dementia was defined according to DSM–III–R criteria (American Psychiatric Association, 1987) and required evidence of neuropsychological deficit and impairment in social or occupational function as evidenced by the formal functional assessments, elicited history, or both. When dementia was diagnosed, all available data were evaluated to determine the type of dementia present. The criteria of the NINCDS-ADRDA were used for the diagnosis of probable or possible AD or other dementias, and CDR scores were applied.

Possible and probable AD subjects included those with sufficient functional and cognitive impairment. Subjects diagnosed with non-Alzheimer dementia (e.g., frontal lobe dementia, Creutzfeldt-Jakob disease, Huntington's disease) were excluded ($n = 506$), as were subjects diagnosed with probable AD and any other concomitant disease (e.g., stroke, Parkinson's disease; $n = 198$).

AD patients were divided into subgroups of disease severity based on their scores on the modified MMSE (mMMSE; range of 0–57). Specifically, 277 subjects with mMMSE scores between 45 and 36 were classified as mild AD. Because we performed stratified random sampling to select a subsample of non-demented elderly controls and mild AD subjects who were matched on age and education (described in detail in the Statistics section), the total number of mild AD subjects was reduced to 244. AD subjects with MMSE scores between 35 and 26 were classified as *moderate* ($n = 480$), between 25 and 16 as *severe* ($n = 376$), and between 15 and 11 as *very severe* AD ($n = 247$). It is important to note that the AD subgroups were labeled as mild, moderate, severe and very severe solely for the purposes of this study; therefore, they reflect mental status levels only and do not necessarily correspond with other such descriptors used in previous studies.

A diagnosis of QD was determined by consensus if a patient scored 45 or greater on the mMMSE, was assigned a CDR score of 0.5 by the examining neurologist, and fulfilled one of the following criteria: (1) insufficient cognitive impairment for a diagnosis of dementia; (2) sufficient cognitive impairment, but insufficient functional impairment for dementia. The diagnosis of QD was based on cognitive impairment across all domains and thus not limited to memory. For the purpose of this analysis, subjects assigned a CDR score of 0.5 but with significant functional impairment (defined as a positive score on the basic ADL section of the Blessed) were excluded ($n = 53$). A positive score on the instrumental ADL section of the Blessed was not considered to reflect significant impairment, and therefore subjects who endorsed any of the questions from this section of the questionnaire were included in the QD group. Subjects diagnosed with QD whose cognitive difficulty was attributed to stroke ($n = 9$), depression ($n = 21$), psychosis or other psychiatric syndrome ($n = 3$), epilepsy ($n = 2$), Parkinson's disease ($n = 3$), alcohol ($n = 3$), drug abuse ($n = 2$), or head trauma ($n = 1$) were also excluded from the current analysis.

Subjects judged to have neither dementia nor QD were classified as non-demented elderly controls ($n = 230$). Again, due to stratified random sampling performed within the non-demented elderly control and mild AD groups, the final QD group was reduced ($n = 155$).

All control subjects were age 55 and older, were assigned a score of zero on the CDR scale, and had mMMSE scores greater than 45. Subjects determined to have depression, psychiatric syndrome, or substance abuse were excluded from this group ($n = 8$). This group, as well as the QD group, includes some individuals who came to medical attention because they or others perceived a problem, but this is representative of non-demented elders who present for diagnostic neuropsychological evaluation.

Neuropsychological Evaluation

The neuropsychological battery used in the present study was comprised of measures used to assess cognitive functions typically affected in dementia and included tests of verbal and nonverbal memory, language, attention, visuospatial ability and abstract reasoning. Written informed consent was obtained from all subjects prior to testing. All subjects were administered a "core" battery, the development of which has been described (Stern et al., 1992). A subset of subjects with mini mental scores over 35 were also administered additional tests, referred to as a "long" battery. In general, patients scoring below 35 on the mMMSE do not require these additional tests to assess the presence of dementia; likewise, they are often unable to complete such a lengthy, complex test session. Tests from the long battery include the Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler, 1981) and selected subtests from the Wechsler Memory Scale–Revised (WMS–R; Wechsler, 1987). Specific analyses were performed on a subset of tests that were selected from both the core and long batteries to represent specific cognitive domains. However, all test scores are presented for the purposes of providing raw score means for a large sample of subjects on various neuropsychological tests. The tests used in the analyses are described below:

Memory

The *Selective Reminding Test* (SRT; Buschke & Fuld, 1974) was used to assess word list learning and verbal memory. Subjects were given six trials to learn a list of 12 unrelated words. After each attempt at recalling the list, the subject was reminded only of the words that were not recalled and then asked to attempt to again to recall the entire list. The number of words learned over six trials, or the total recall score, was used to assess short-term verbal memory. Sub-

jects were asked to recall the original list following a 15-min delay. The number of words recalled at this time was used to determine delayed recall. After completing the SRT, recognition of words that were not recalled was then tested with the use of multiple-choice arrays. This score was used as a measure of delayed recognition.

The *Logical Memory I and II* subtests of the WMS–R (Wechsler, 1987) were used to assess verbal memory as well. For this test subjects listened to two short stories after which their recall was tested (immediate recall) and again following a delay of 15 min (delayed recall).

The *Visual Reproduction I and II* subtests of the WMS–R were used as measures of nonverbal memory. Subjects were shown four designs and asked to draw each from memory (immediate recall). Following a 15-min delay the subject was instructed to again reproduce each design (delayed recall).

A multiple-choice version of the *Benton Visual Retention Test* (BVRT; Benton, 1955) was used to assess nonverbal memory in a format that did not rely on constructional abilities. The subject viewed a design for 10 s. The design was then removed and the subject was asked to recognize the design in an array that included three distracters. Ten stimulus items were used, corresponding to Form D of the original Benton Visual Retention Test.

Language

On the *Boston Naming Test* (BNT; Kaplan, 1983), a test of word-finding ability, the subject was presented with 60 line drawings of objects and instructed to give the objects' names. If unable to name an object after 20 s, a semantic hint was offered by the tester. If still unable, the subject was then provided a phonemic clue. A 15-item version of this test was administered to patients with mMMSE scores less than 35 (the results of this version are found in Table 3), but scores from the full BNT were used for the present study.

The *Controlled Oral Word Association* test from the Multilingual Aphasia Examination (Benton & Hamsher, 1976) was used to assess verbal fluency. Here, the subject was given one minute each to name as many words as possible, beginning with the letters *C*, *F*, and *L*. The mean of the raw scores on all three trials was used.

On the *Category Naming* test the subject was instructed to name as many items within a specified category as possible. One minute was allowed for each of the three categories: animals, food, and clothing. Scores were expressed in terms of the mean of the raw scores on all three trials.

The high frequency items from the *Repetition of Phrases* subtest of the Boston Diagnostic Aphasia Evaluation (BDAE; Goodglass, 1983) were used. On this test, subjects are asked to repeat phrases read by the examiner. The total number of correctly repeated phrases was used.

Attention

The *Digit Span, Digit Symbol, and Arithmetic* subtests of the WAIS–R were used as measures of freedom from distraction. Raw scores from all of these subtests were used, and the sum of the Digit Span forward and reverse raw scores were used to represent the Digit Span score.

Two timed paper and pencil tests, or *Cancellation Tests*, were used. For each, a single page with a target stimuli appearing at the top followed by rows and columns of the stimuli was administered. One form of the test (shape) featured a diamond shaped stimulus, which the subject was instructed to cross out as quickly as possible. In the second form, the target was a group of letters (*TMX*). While speed was recorded (the results of which are presented in Table 3), for this study performance was measured as the number of items incorrectly omitted for each form of the test (omits).

Visuospatial ability

The *Block Design* and *Object Assembly* subtests of the WAIS–R were used. The Block Design test requires the subject to reproduce a series of designs using red and white colored blocks within a time limit, while the Object Assembly test involves placing pieces together, as in a puzzle. Raw scores on both tests were used in this study.

On the *Rosen Drawing Test* (Rosen, 1981), a test of graphomotor construction ability, the subject copied five designs that were selected from the original Rosen Drawing Test to span a range of difficulty from simple shapes and topological concepts to overlapping, Euclidean, and three-dimensional designs. The total score was used for this analysis.

The matching subtest of the *Benton Visual Retention Test* was used. For each of ten items, the subject matched a larger picture to one in an array of four smaller pictures. Items corresponded to Form C of the original Benton Visual Retention Test.

Abstract reasoning

A subset of questions from the *Similarities* subtest of the WAIS–R were used. This test required the subject to identify relevant similarities between pairs of items. Raw scores were used.

For the *Identities and Oddities* subtest of the Mattis Dementia Rating Scale (Mattis, 1976), the subject was asked to identify which two of three items were the same. After eight trials were completed, the same items were administered again, with the subject identifying the one item that was different. The total number of correct items were combined to determine the score.

Statistics

Stratified random sampling was applied to the elderly control and mild AD groups to select a sample of subjects who were matched in age and education. As a result, the size of the control group was reduced from 230 to 155, and the mild AD group was decreased from 277 to 244.

Two general sets of analyses were performed. In the first, descriptive data from all six groups was examined. One way analyses of variance and Tukey HSD *post-hoc* tests were used to compare the continuous demographic vari-

ables, functional scores, and mean scores on each test from both the core and long battery across all six groups. Because the long battery was not administered to the severe and very severe subjects, these groups were not included in the ANOVAs run on the long battery data. Chi-square analyses and Duncan's multiple range *post-hoc* comparisons were used to assess the gender groupings of the subjects.

To examine the possibility of pre-existing cognitive differences between groups, estimates of premorbid IQ were obtained for control subjects by applying the Barona equation, a demographically based index of premorbid intelligence (Barona et al., 1984; see Appendix A for exact regression equation). One way analyses of variance were used to compare the elderly controls' estimated verbal IQ scores to their WAIS–R verbal IQ scores, and also to compare the Barona estimated scores between groups.

For the second set of analyses, data from three of the six original groups was used: elderly controls, QD subjects, and mild AD subjects. Individual neuropsychological test scores of the QD and mild AD groups were transformed to create *z* scores using means and standard deviations of the elderly control subjects based on the procedure described by (Huff, 1987). The purpose of creating *z* scores was twofold: it permitted us to compare performance across groups on individual tests within cognitive domains and to form composite scores for each domain. Composite scores were created by averaging the *z* scores of individual tests selected as representative of each domain (memory, language, attention, visuospatial ability, and abstract reasoning), as performed in previous examinations of the cognitive deficits of AD (Becker et al., 1994; Huff et al., 1987).

Two sets of analyses were conducted to predict the categorical outcome of group membership. First, the five composite domain scores were used as covariates in two separate logistic regressions to determine which domains best differentiated between (1) elderly control subjects and QD subjects, and (2) elderly control subjects and mild AD subjects. Age and education were included as additional covariates and the variables were entered into the analyses using a forward stepwise procedure.

For the second set of analyses, two separate logistic regressions were conducted on the individual neuropsychological test data within each domain to determine those most predictive of group status. In light of the fact that individual tests often make demands on a variety of cognitive skills, we included all of the tests used to create the 5 domains, regardless of whether a particular domain remained in the initial logistic regression equations. Therefore, for each of the five domains, two stepwise forward logistic regressions were used, with group as the dependent variables (control *vs.* QD, control *vs.* mild AD). The mean *z*-scores obtained by each group on those tests within the domain, as well as age and education, were entered as the covariates.

RESULTS

Demographic values, mean mMMSE scores, CDR scores, and mean functional ADL scores are presented in Table 1. Because we used stratified random sampling, the three groups of primary interest to our analyses did not differ in age. The three remaining groups of AD patients were significantly older. Education differences were found, with education decreasing in the moderate, severe, and very severe groups. Again, our sampling procedure resulted in equal levels of education for the elderly controls and mild AD patients. QD patients, who were not involved in our sampling procedure, had similar levels of education to both elderly controls and mild AD subjects. Mental status scores and functional scores reflect our classification criteria and are included only to illustrate this process.

The mean neuropsychological scores from the core battery are presented in Table 2, and the scores from the long battery can be found in Table 3. These scores are provided for descriptive purposes and provide a reference for the pattern of mean scores seen on this comprehensive battery. Results of the univariate ANOVAs on the mean core and

Table 1. Demographic features and mean mMMSE, CDR and ADL scores of non-demented elderly controls, QD and AD groups

Variable	Range	Non-demented elderly controls (N = 155)	Questionable dementia (N = 100)	Mild AD (mMMS>45) (N = 244)	Moderate AD (mMMS>35) (N = 480)	Severe AD (mMMS>25) (N = 376)	Very severe AD (mMMS>10) (N = 247)
Age	56–90	67.9 ± 8.2[a]	69.7 ± 10.5[a]	71.6 ± 9.6[a]	73.4 ± 9.1[b]	74.6 ± 9.1[b]	74.6 ± 9.2[b]
Years of Education	0–20	15.1 ± 3.5[ab]	14.3 ± 3.9[bc]	15.6 ± 3.5[a]	13.2 ± 3.8[c]	11.8 ± 4.1[d]	10.6 ± 5.2[e]
Gender		40% Male 60% Female	49% Male 51% Female	55% Male 45% Female	38% Male 62% Female	31% Male 69% Female	33% Male 67% Female
Mean mMMSE	0–57	52.4 ± 3.3[a]	50.3 ± 3.0[b]	48.7 ± 2.8[c]	39.6 ± 2.7[d]	29.5 ± 2.8[e]	18.5 ± 4.0[f]
Mean CDR score	0–5	0.0[a]	0.5[b]	1.0 ± .90[c]	1.0 ± 2.0[c]	1.2 ± 5.0[d]	1.6 ± 6.5[e]
Mean Blessed FAS	0–17	.95 ± .86[a]	1.2 ± .99[a]	2.3 ± 1.7[b]	3.0 ± 2.0[c]	3.9 ± 2.6[d]	5.8 ± 3.3[e]
Mean Blessed Basic ADL	0–9	0.0[a]	0.0[a]	.28 ± .86[a b]	.46 ± .97[b]	.82 ± 1.4[c]	1.6 ± 1.9[d]
Mean ADL % (rated by examiner)	0–100	92.7 ± 10.3[a]	91.1 ± 9.7[a]	83.6 ± 13.5[b]	77.9 ± 17.8[c]	68.4 ± 20.4[d]	56.2 ± 22.4[e]

Analyses reflect differences at the .001 level for comparison of adjacent groups. Values with different superscript letters differ significantly on *post-hoc* testing. Values are means ± standard deviations. Values in brackets represent number of subjects.

Table 2. Mean scores and standard deviations for clinical groups on core battery neuropsychological tests

Measure	Range	Non-demented elderly mMMSE>45	Questionable dementia mMMSE>45	Mild AD mMMS>45	Moderate AD mMMS>35	Severe AD mMMS>25	Very severe AD mMMS>10
SRT Total Recall	0–72	42.9 (9.8)[a] [153]	35.7 (10.5)[b] [94]	28.7 (8.8)[c] [221]	23.6 (8.5)[d] [452]	18.0 (7.9)[e] [336]	11.4 (7.2)[f] [153]
SRT Delay Recall	0–12	6.2 (3.0)[a] [153]	3.8 (2.4)[b] [94]	1.9 (2.0)[c] [219]	1.3 (1.8)[d] [450]	0.71 (1.3)[e] [331]	0.28 (0.7)[e] [147]
SRT Delay Recognition	0–12	11.3 (1.2)[a] [153]	10.5 (1.7)[b] [94]	9.0 (2.7)[c] [219]	7.7 (2.8)[d] [446]	6.0 (2.6)[e] [324]	4.7 (2.6)[f] [141]
BVRT Recognition	0-10	8.2 (1.7)[a] [152]	7.6 (1.4)[a b] [91]	7.1 (2.0)[b] [219]	5.7 (2.1)[c] [419]	4.4 (2.0)[d] [309]	3.5 (1.6)[e] [129]
BVRT Matching	0-10	9.4 (0.9)[a] [152]	9.4 (0.9)[a] [92]	9.0 (1.2)[a] [219]	7.9 (2.0)[b] [423]	6.6 (2.3)[c] [314]	5.2 (2.5)[d] [130]
Rosen Drawing (5 item)	0-5	3.3 (1.1)[a] [155]	3.1 (0.9)[a] [92]	3.0 (1.1)[a] [220]	2.6 (1.1)[b] [442]	2.0 (1.3)[c] [343]	1.3 (1.1)[d] [167]
BNT (15 item)	0-15	14.1 (1.3)[a] [153]	13.9 (1.8)[a] [92]	13.5 (2.0)[a] [212]	12.0 (2.6)[b] [427]	10.3 (3.0)[c] [347]	7.4 (3.1)[d] [167]
CFL Mean	N/A	13.0 (4.9)[a] [154]	12.3 (5.1)[a b] [94]	11.3 (4.9)[b] [220]	8.2 (4.3)[c] [443]	5.6 (3.4)[d] [336]	3.3 (2.3)[e] [153]
Category Fluency Mean	N/A	18.6 (8.3)[a] [147]	15.6 (4.2)[b] [93]	13.2 (5.1)[c] [219]	9.7 (4.0)[d] [427]	7.1 (2.7)[e] [330]	4.4 (2.7)[f] [154]
Repetition	0–8	7.7 (0.8)[a] [152]	7.6 (0.6)[a] [90]	7.8 (0.5)[a] [214]	7.5 (0.9)[a] [442]	7.1 (1.3)[b] [349]	6.1 (1.8)[c] [169]
Similarities	0–28	17.7 (6.1)[a] [151]	17.2 (5.8)[a b] [83]	15.9 (5.9)[b] [204]	9.8 (6.1)[c] [416]	5.4 (5.0)[d] [345]	3.0 (3.8)[e] [170]
Cancellation (Shape time in sec)	20–240	63.1 (27.0)[a] [150]	63.3 (22.9)[b] [91]	77.8 (39.3)[c] [215]	103.4 (50.1)[d] [397]	139.5 (60.9)[d] [279]	159.7 (58.2)[d] [99]
Cancellation (Shape omits)	0–20	4.3 (3.7)[a] [150]	4.9 (3.7)[a b] [91]	6.2 (4.3)[b c] [215]	7.1 (4.7)[c] [397]	8.7 (5.3)[d] [279]	9.0 (5.4)[d] [99]
Cancellation (TMX time in sec)	20–240	71.9 (26.0)[a] [151]	75.5 (23.4)[a] [91]	81.8 (33.6)[a] [216]	107.2 (45.9)[b] [403]	144.1 (59.1)[c] [286]	169.8 (60.5)[d] [107]
Cancellation (TMX omits)	0–20	1.0 (1.6)[a] [151]	1.3 (2.5)[a b] [91]	1.7 (2.5)[a b] [215]	2.3 (3.0)[b] [403]	4.9 (4.8)[c] [286]	6.6 (5.0)[d] [107]
Comprehension	0–6	5.6 (0.9)[a] [152]	5.7 (0.7)[a] [90]	5.4 (0.9)[a] [209]	4.8 (1.2)[b] [436]	3.9 (1.5)[c] [346]	2.7 (1.6)[d] [165]
Identities/Oddities	0–16	15.1 (1.7)[ab] [147]	15.2 (1.4)[a] [90]	14.5 (1.9)[b] [214]	13.5 (2.3)[c] [411]	12.6 (2.4)[d] [325]	11.1 (3.0)[e] [147]

Values are means ± standard deviations. Values in brackets represent number of subjects.
Analyses reflect differences at the .01 level for comparison of adjacent groups. Values with different superscript letters differ significantly on *post-hoc* testing.

long battery scores illustrate that QD subjects performed no differently from normal controls on some tests, and similar to mild AD patients on others.

The elderly controls' estimated verbal IQ scores did not differ from their actual WAIS–R VIQ scores (114.7 ± 7.3 *vs.* 115.0 ± 12.5, respectively), thus we can assume that the Barona equation provides a valid estimate of premorbid cognitive functioning. Based on this assumption, results of the between-group analysis on WAIS–R VIQ scores reveal that the premorbid functioning of the elderly control subjects was significantly higher than that of the moderate AD group ($p < .01$), (but not the QD and mild AD group). See Table 4 for comparison of Barona estimated VIQ scores across groups.

The *z* scores constructed for our logistic regression analyses can be found in Table 5. For the analysis comparing the neuropsychological scores of the control subjects and QD subjects, Table 6 shows that only the memory domain contributed to the regression model (OR = 2.8, CI = 1.5–5.0). This odds ratio indicates that after adjustment for other variables in the model, for every unit decrease in the memory domain score, the odds of being in the QD group are increased almost three-fold.

Results of the analysis comparing control subjects and mild AD patients revealed that the memory domain provided the greatest contribution to the model (OR = 11.0, CI = 4.9–25.1), followed by the abstract reasoning domain (OR = 4.1, CI = 1.3–12.6). These odds ratios reveal that for every unit decrease in the memory domain score, a subject is 11 times as likely to be in the mild AD group, and for every unit decrease in the abstract reasoning domain, these odds are increased 4 times.

Table 7 contains the odds ratios and confidence intervals for the logistic regressions conducted on individual neuro-

Table 3. Mean scores and standard deviations for clinical groups on full battery neuropsychological tests

Measure	Range	EC	QD	Mild AD	Moderate AD	Severe AD	Very severe AD
WAIS–R subtests							
Full Scale IQ	50–150	113.0 (13.5)[a] [77]	106.1 (12.9)[b] [71]	102.3 (12.7)[b] [166]	86.7 (9.5)[c] [230]	78.5 (9.4) [23]	82.0 (3.4) [3]
Verbal IQ	50–150	115.0 (12.5)[a] [77]	109.3 (12.7)[b] [71]	105.8 (12.3)[b] [166]	91.0 (10.0)[c] [239]	82.8 (9.2) [24]	85.3 (5.6) [3]
Perfor IQ	50–150	106.7 (14.9)[a] [77]	100.3 (14.2)[b] [71]	95.4 (13.7)[b] [166]	82.3 (11.4)[c] [237]	75.6 (14.2) [24]	78.0 (6.0) [3]
Information (raw score)	0–29	23.1 (4.3)[a] [77]	21.4 (4.6)[ab] [67]	19.5 (5.5)[ab] [171]	14.0 (5.4)[c] [251]	6.2 (5.9) [38]	5.8 (8.5) [7]
Vocabulary (raw score)	0–70	58.5 (8.6)[a] [77]	52.2 (12.0)[b] [68]	51.4 (12.4)[b] [169]	39.3 (15.2)[c] [252]	21.9 (19.5) [37]	17.0 (21.4) [7]
Arithmetic (raw score)	0–19	12.6 (3.7)[a] [76]	11.8 (3.6)[ab] [66]	10.7 (4.0)[b] [167]	6.9 (3.0)[c] [241]	3.7 (3.3) [36]	2.5 (3.4) [7]
Comprehen (raw score)	0–32	24.2 (4.5)[a] [76]	22.8 (4.5)[ab] [66]	19.7 (6.7)[b] [159]	13.9 (6.3)[c] [235]	8.2 (8.0) [36]	2.8 (4.1) [7]
PictureCom (raw score)	0–20	14.3 (3.8)[a] [77]	12.7 (4.1)[ab] [67]	11.3 (4.6)[b] [168]	6.9 (4.3)[c] [256]	3.0 (4.0) [39]	3.4 (4.7) [7]
Pic Arrang (raw score)	0–20	10.7 (4.5)[a] [76]	8.3 (4.2)[b] [67]	6.2 (4.2)[c] [164]	3.5 (2.8)[d] [239]	1.4 (2.3) [35]	.71 (.95) [7]
Digit Span (raw score)	0–28	16.0 (3.9)[a] [77]	14.7 (4.3)[ab] [67]	14.2 (3.9)[b] [169]	11.5 (3.9)[c] [258]	6.7 (5.4) [39]	4.1 (5.2) [7]
Digit Span Forward	0–9	6.8 (1.3)[a] [78]	6.5 (1.3)[a] [69]	6.3 (1.4)[a] [170]	5.7 (1.5)[b] [258]	3.6 (2.7) [38]	2.2 (2.8) [7]
Digit Span Reverse	0–8	5.2 (1.2)[a] [78]	4.6 (1.2)[b] [69]	4.5 (1.2)[b] [170]	3.8 (1.2)[c] [257]	2.2 (1.7) [38]	1.2 (1.6) [7]
Block Des (raw score)	0–51	24.5 (9.1)[a] [77]	19.5 (10.3)[b] [67]	16.4 (9.3)[b] [168]	8.9 (7.7)[c] [248]	4.7 (8.2) [38]	1.3 (2.8) [8]
Object Ass (raw score)	0–41	24.5 (8.9)[a] [75]	21.5 (7.5)[ab] [67]	18.6 (8.1)[b] [166]	14.0 (8.0)[c] [238]	7.1 (7.8) [37]	6.7 (8.1) [8]
Digit Symb (raw score)	0–93	41.2 (13.9)[a] [77]	36.5 (10.0)[ab] [67]	32.1 (12.2)[b] [167]	21.1 (11.6)[c] [241]	11.7 (14.2) [37]	6.0 (9.2) [7]
WMS–R subtests							
Logical Memory	0–48	8.9 (2.6)[a] [76]	7.5 (4.0)[b] [65]	5.3 (2.6)[c] [164]	3.7 (3.0)[d] [233]	1.0 (1.5) [38]	0.4 (0.6) [7]
Log Mem Delay	0–48	6.9 (2.9)[a] [76]	5.1 (3.6)[b] [65]	2.7 (2.6)[c] [163]	1.2 (2.0)[d] [226]	0.6 (1.8) [34]	0.1 (0.2) [7]
Visual Reproduc	0–14	7.8 (3.2)[a] [76]	6.3 (3.4)[b] [65]	4.7 (2.9)[c] [161]	2.5 (2.6)[c] [234]	1.6 (2.6)[d] [36]	0.7 (1.2) [7]
Vis Repro Delay	0–14	6.0 (3.6)[a] [75]	3.5 (3.4)[b] [65]	1.7 (2.1)[c] [161]	0.6 (1.4)[d] [228]	0.7 (2.4) [34]	0.2 (0.7) [7]
Vis Repro Recog	0–4	3.2 (1.0)[a] [63]	2.7 (0.9)[ab] [57]	2.4 (1.7)[b] [124]	1.7 (1.2)[c] [169]	0.8 (1.1) [28]	0.6 (1.2) [6]
Paired Assoc	0–21	13.9 (4.0)[a] [75]	11.9 (3.2)[b] [64]	9.2 (3.6)[c] [163]	7.6 (3.2)[d] [219]	3.5 (3.4) [37]	3.0 (3.8) [7]
Paired Ass. Del	0–7	5.5 (1.5)[a] [75]	4.5 (1.5)[ab] [64]	3.3 (1.6)[c] [156]	2.8 (1.5)[c] [205]	1.6 (2.0) [35]	1.4 (1.9) [7]
Additional tests							
BNT (60-item)	0–60	52.1 (10.0)[a] [80]	49.3 (10.9)[ab] [76]	45.7 (12.3)[b] [182]	37.5 (14.0)[c] [265]	18.0 (16.2) [42]	19.2 (26.0) [7]
Rosen 19-item	0–19	15.3 (3.1)[a] [80]	14.0 (2.7)[ab] [76]	13.4 (3.4)[b] [180]	11.0 (3.8)[c] [279]	7.8 (6.1) [43]	4.5 (5.8) [7]

Values are means ± standard deviations. Values in brackets represent number of subjects.
Analyses reflect differences at the .01 level for comparison of four of the six groups (due to small *n* sizes in the severe and very severe groups, they were excluded from analyses) . Values with different superscript letters differ significantly on *post-hoc* testing.

psychological scores comprising each domain. Within the memory domain, four of the original test scores were retained in the model comparing control subjects and mild AD subjects: SRT total recall, SRT delayed recall, Logical Memory immediate recall, and Visual Reproduction delayed recall. When control subjects were compared to QD subjects, only one test, SRT delayed recall, was effective at separating the groups. Within the language domain, the cat-

Table 4. Comparison of Barona's estimated IQ scores across groups

	Elderly control ($N = 77$)	Questionable dementia ($N = 71$)	Mild AD ($N = 166$)	Moderate AD ($N = 239$)
Barona estimated Verbal IQ	114.7 (7.3)*	110.7 (9.1)	113.5 (7.2)	107.5 (10.5)*

*$p < .01$
Values are means ± standard deviations.

egory fluency score was the only test to be retained in both models. When the attention domain was analyzed, the Digit Symbol, Digit Span, and Cancellation (shape omits) tests contributed to the model that differentiated between control and mild AD subjects, while only Digit Symbol separated the control and QD subjects. Within the visuospatial domain, Block Design was the only test score to contribute to the model differentiating control and mild AD subjects, while none were retained in the model separating control subjects and QD subjects. Finally, the comparison between controls and mild AD subjects for the abstract reasoning domain revealed that both the Similarities and Identities and Oddities tests were effective at predicting group membership, while no tests were retained when comparing controls and QD subjects.

Table 5. Normalized *z*-scores and standard errors for QD and mild AD groups (based on mean and standard deviation of control groups) on neuropsychological tests used in logistic regressions

Cognitive domain	Questionable dementia	Mild AD
Memory		
SRT Total Recall	−.57 (.13) [71]	−1.3 (.06) [164]
SRT Delay Recall	−.73 (.09) [71]	−1.3 (.05) [163]
SRT Delay Recognition	−.51 (.16) [71]	−1.6 (.15) [163]
Benton Recognition	−.24 (.10) [68]	−.45 (.08) [163]
WMS–R Logical Memory	−.53 (.18) [65]	−1.3 (.07) [159]
WMS–R Log Mem Delay	−.60 (.15) [65]	−1.4 (.07) [158]
WMS–R Vis Reproduction	−.45 (.13) [65]	−.92 (.07) [156]
WMS–R Vis Repro Delay	−.69 (.11) [65]	−1.1 (.04) [156]
Language		
Boston Naming Test	−.28 (.13) [69]	−.56 (.08) [163]
CFL Mean	−.03 (.12) [71]	−.26 (.08) [165]
Category Fluency Mean	−.24 (.05) [70]	−.58 (.04) [165]
Repetition	−.05 (.09) [67]	−.05 (.04) [160]
Attention		
Digit Span Raw	−.31 (.13) [67]	−.39 (.07) [165]
Arithmetic Raw	−.20 (.12) [66]	−.50 (.08) [167]
Shape Omits	−.07 (.09) [68]	−.51 (.09) [162]
Digit Symbol Raw	−.33 (.08) [67]	−.65 (.06) [167]
TMX Omits	−.06 (.16) [68]	−.22 (.09) [161]
Visuospatial functioning		
Block Design Raw	−.54 (.13) [67]	−.88 (.07) [165]
Rosen Drawing Test	−.13 (.08) [69]	−.11 (.07) [162]
Benton Matching	.22 (.07) [69]	−.14 (.07) [163]
Object Assembly Raw	−.33 (.08) [67]	−.65 (.06) [167]
Abstract reasoning		
Similarities Raw	.13 (.10) [67]	−.14 (.06) [165]
Identities/Oddities	.12 (.07) [67]	−.23 (.07) [162]

Values are means ± standard errors. Values in brackets represent number of subjects.

DISCUSSION

The intention of this study was to compare, cross-sectionally, the neuropsychological scores of a very large clinically based sample that included AD patients of varying severity, QD subjects, and non-demented elderly controls. Neuropsychological functioning was assessed by means of a test battery typically administered in specific dementia evaluations that provided a standard data collection procedure for obtaining scores for all subjects on the same variables (i.e., "core battery"). Additional measures were administered when appropriate, given the mental status of the examinee, and included tests used in most neuropsychological evaluations (e.g., WAIS–R, WMS–R). All tests were assembled so as to best measure functioning across verbal and non-verbal, and auditory and visual modalities. One of the primary goals of the study, given its large scope, was to provide, in an accessible form, mean scores for each group of subjects with the assumption that other clinicians could refer to these scores when examining older patients for whom memory impairment is a presenting problem.

A second goal of this study was to compare the neuropsychological performance of subjects with QD (who are not demented but demonstrate cognitive impairment beyond that expected for age and education) and mild AD to that of non-demented elderly controls, whose performance reflects age-related cognitive decline, or "normal aging" effects. This objective was achieved by performing logistic regressions on cognitive domains and individual tests within

Table 6. Forward stepwise logistic regression results comparing QD and mild AD groups with elderly controls (EC) on all domains

Comparison	Domain	Odds ratio	CI
EC *vs.* QD ($n = 255$)	Memory	2.8	1.50–5.0
EC *vs.* mild AD ($n = 399$)	Memory	11.0	4.8–25.1
	Abstract reasoning	4.1	1.3–12.6

Table 7. Forward stepwise logistic regression results comparing QD and mild AD groups with elderly controls on individual test scores within domains

Comparison	Domain	Test	Odds ratio	CI
EC *vs.* QD	Memory	SRT Delayed Recall	2.2	1.5–3.2
	Language	Category Fluency	4.2	1.9–9.1
	Attention	Digit Symbol	1.7	1.0–2.7
EC *vs.* mild AD	Memory	SRT Total Recall	1.8	1.0–3.4
		SRT Delayed Recall	2.1	1.1–3.8
		Logical Memory I	2.3	1.4–3.8
		Visual Repro II	2.5	1.4–4.4
	Language	Category Fluency	8.9	4.7–17.0
	Attention	Digit Symbol	2.4	1.6–3.6
		Cancellation (omits)	1.7	1.2–2.4
		Digit Span	1.4	1.0–1.9
	Visuospatial	Block Design	2.3	1.6–3.2
	Abstract Reasoning	Similarities	1.4	1.1–1.9
		Identities/Oddities	1.3	1.0–1.8

domains in order to identify a profile of performance associated with group membership.

Elderly Controls and QD Subjects

Our results illustrate that compared to that of non-demented elderly individuals, the memory performance of QD subjects was most predictive of group membership. An analysis of specific memory measures revealed that delayed recall of a list of words best differentiated QD subjects from controls, suggesting that acquisition of verbal information, both in list form and in meaningfully organized story form, is relatively intact. Recognition performance and recall of nonverbal information also did not differentiate the groups, leading us to conclude that the lower memory performance in QD patients, as compared to non-demented elderly controls, is restricted to the retrieval, or delayed recall of verbal information organized in list form.

While our regression model predicting membership into the QD or elderly control groups did not retain any domains other than memory, our regressions using individual test comparisons revealed additional predictive tests *within* the language and attention domains. Category fluency was found to contribute to the model predicting QD or elderly control membership. The fact that phonemic fluency performance was not retained in the model illustrates the higher sensitivity of category fluency measures as compared to letter fluency due to the semantic knowledge required for the former (Butters et al., 1987). This difference has been demonstrated repeatedly in AD samples and has been shown to discriminate between normal controls and mild AD patients (Monsch et al., 1992).

Membership in the QD group was also predicted, in part, by the Digit Symbol test, a measure requiring the subject to "attend to, process, and remember" (Wechsler, 1981). Other attention measures were not retained, however (i.e., Digit Span, Arithmetic), suggesting that the additional processing component involved in the Digit Symbol test, as well as the psychomotor speed component, may have contributed to the lower performance of the QD group. When we consider that decreased psychomotor speed is associated with AD (Storandt & Hill, 1989) and worsens with the severity of dementia, our finding of decreased Digit Symbol performance for the QD subjects suggests that this measure may provide a sensitive means of differentiating QD from normal aging. Overall, these results support previous findings in regard to the cognitive deficits characterized in QD, such as delayed recall (Petersen et al., 1999). They are also consistent with structural abnormalities identified among QD and MCI patients with abnormal memory, including decreased volume in the hippocampus (Wolf et al., 2001), parahippocampal gyrus, entorhinal cortex, and superior temporal gyrus (Killiany et al., 2000; Visser et al., 1999). The presence of the category fluency and Digit Symbol tests in the regression models represent areas rarely reported in studies of QD; nonetheless, our results are not surprising given reports of impairments in these areas in previous studies of AD (Masur et al., 1990; Monsch et al., 1992). We can conclude, then, that subtle deficits may arise in these areas within QD subjects.

Elderly Controls and Mild AD Subjects

While two domains, memory and abstract reasoning, best differentiated the elderly control and mild AD groups, further analyses revealed reduced performance for the mild AD subjects on tests within each of the five domains. Deficient delayed recall of list-organized material and immediate recall of stories (representing meaningfully organized material) distinguished the mild AD group within the memory domain, consistent with the well-established finding of decreased ability to learn new information in AD (Morris et al., 1991; Welsh et al., 1991). The abstract reasoning domain provided the second greatest additional contribu-

tion to the equation for group membership. Both the verbal (Similarities) and nonverbal (Identities and Oddities) measures used in this domain were lower in the mild AD group, providing a robust illustration of this impairment early in the disease.

Performance on category fluency, the same language test that differentiated the QD subjects from elderly controls, also differentiated the mild AD group from controls. As expected based on previous findings (e.g., Monsch et al., 1992), letter fluency performance was not retained in the model. Within the attention domain, the mild AD subjects demonstrated lower performance on three of the four measures, including the Digit Symbol test, which was discussed earlier in the comparison of QD and elderly control subjects. However, for the mild AD subjects, auditory attention was also affected (Digit Span), as was visual scanning ability (Cancellation). These findings are in line with longitudinal studies of mild AD patients who demonstrate deficits on tests with attention components (Haxby et al., 1988). In the visuospatial domain, the only test that differentiated mild AD patients from elderly controls was a measure requiring visual organization, problem solving and psychomotor speed (Block Design). Since this test involves such a variety of components, it is difficult to determine which specific abilities may be compromised.

As expected, our study demonstrates greater overall cognitive impairment in mild AD compared to QD or normal aging, with a specific deficit in free recall of recently acquired information. These results are in line with pathological findings in the early stages of AD, when medial temporal limbic areas, particularly hippocampal atrophy and entorhinal cortex volume, are affected; atrophy in neocortex areas is typically associated with increased disease severity (Jack et al., 2002). Although it is clear that cognitive decline in AD is progressive, the heterogeneity of deficits and variability between patients is well known (Mayeux et al., 1985). Measuring cognitive functioning at one point in time did not allow us to predict the rates of progression for impairment. Nonetheless, we have the ability, due to our large sample size and comprehensive test battery, to provide descriptive profiles of those tests found to be most helpful for differentiating between QD and mild AD without the confounding effects of normal aging. Specifically, AD subjects demonstrated deficient learning of both semantically organized (Logical Memory I) and list-organized verbal material (SRT Total Recall), as well as deficient delayed recall of verbal (SRT Delayed Recall) and nonverbal material (Visual Reproduction II). Category fluency was lowered (while letter fluency remained relatively preserved), as was performance on tests of attention (Digit Symbol, Cancellation (number of omits), Digit Span) and visuospatial ability (Block Design). Finally, abstract reasoning was deficient (Similarities and Identities/Oddities). QD subjects demonstrated a subset of these deficiencies when compared to elderly controls, with impairment on delayed recall of list items (SRT delayed recall), Category Fluency, and Digit Symbol (see Table 7). These profiles support previous findings regarding the cognitive deficits associated with mild AD and QD, particularly impaired delayed recall of verbal information. The superiority of category fluency over letter fluency in differentiating early AD patients from healthy elderly (Monsch et al., 1992) was also demonstrated in our study, both for our mild AD and QD subjects. As previously indicated, this result, combined with our finding of reduced performance on the Digit Symbol test in both QD and mild AD subjects, suggests that delayed recall may not be the only function affected in questionable dementia; that is, attention and psychomotor speed may represent additional cognitive areas affected.

The present study includes a number of limitations. Because the clinical criteria for QD are imprecise, the QD subjects in this study represent a heterogeneous group. For the purposes of this study a positive score on the instrumental section of the Blessed was not considered to be reflective of functional impairment, since these items are frequently endorsed by both normal and memory-impaired respondents. For example, on a question from the Blessed that inquires about trouble around the house "like cleaning, doing the laundry, cooking . . . and other chores," 21% of the elderly controls and 27% of the QD subjects responded positively. It is important to point out that within these two subject groups, Blessed interviews were typically conducted with the subject; in contrast, information regarding the functional status of the four AD groups was more often provided by the subject with an informant (e.g., a caregiver) or by an informant alone. Ratings provided by a caregiver or informant may be biased, resulting in under- or over-estimation of patient's functioning (DeBettignies et al., 1993). Since functional status was more often self-reported by elderly control and QD subjects, the possibility that they overestimated their functioning or were unaware of their deficits must be acknowledged in light of the diagnostic criteria used in the present study.

A second issue in regard to group selection requires clarification. The neuropsychological scores that were used to determine diagnosis and thus group membership are the same scores that were used in this study to identify specific domains of impairment across groups. While this process may be considered circular in nature, we feel that this method represents a valid means of examining the data given the goals of the present study.

Because our selection of neuropsychological measures was constrained by the availability of tests from our archival data, we did not include tests of executive functioning in our analyses. Since executive dysfunction is commonly considered in the context of intact function in other domains, we have typically not relied on this domain for the evaluation of patients with Alzheimer's disease. Nonetheless, at least one study has reported that specific executive function tests were impaired in mild AD patients (Lafleche & Albert, 1995). The fact that our battery did not incorporate tests of this particular domain limits our ability to comment on all cognitive abilities, and additional research should assess whether tests other than those used in this study might prove more sensitive to the cognitive functioning differences between normal aging, QD, and mild AD. The present

study provides a framework of neuropsychological scores associated with age-related cognitive decline, QD and mild AD, which can be referenced by clinicians in the evaluation of elderly populations.

ACKNOWLEDGMENTS

This work was supported in part by National Institute of Aging Grant, T32AG00261 (Y.S.), AG08702 (Alzheimer's Disease Center).

REFERENCES

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (Rev. 3rd ed.). Washington, DC: American Psychiatric Press.

Barona, A., Reynolds, C., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS–R. *Journal of Consulting and Clinical Psychology*, *52*, 885–887.

Becker, J.T., Boller, F., Lopez, O.L., Saxton, J., & McGonigle, K.L. (1994). The natural history of Alzheimer's disease: Description of study cohort and accuracy of diagnosis. *Archives of Neurology*, *51*, 585–594.

Benton, A.L. (1955). *The Visual Retention Test*. New York: The Psychological Corporation.

Benton, A. & Hamsher, K. (1976). *Multilingual Aphasia Examination*. Iowa City, IA: University of Iowa.

Blessed, G., Tomlinson, B.E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry*, *114*, 797–811.

Buschke, H. & Fuld, P.A. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*, *24*, 1019–1025.

Butters, N., Granholm, E., Salmon, D.P., Grant, I., & Wolfe J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, *9*, 479–497.

Cahn, D.E., Salmon, D.P., Butters, N., Wiederholt, W.C., Corey-Bloom, J., Edelstein, S.L., & Barrett-Connor, E. (1995). Detection of dementia of the Alzheimer type in a population-based sample: Neuropsychological tets performance. *Journal of the International Neuropsychological Society*, *1*, 252–260.

Celsis, P. (2000). Age-related cognitive decline, mild cognitive-impairment or preclinical Alzheimer's disease? *Annals of Medicine*, *32*, 6–14.

Collie, A., Shafiz-Antonacci, R., Maruff, P., Tyler, P., & Currie, J. (1999). Norms and the effects of demographic variables on a neuropsychological battery for use in healthy ageing Australian populations. *Australian and New Zealand Journal of Psychiatry*, *33*, 568–575.

Connor, D.J., Salmon, D.P., Sandy, T.J., Galasko, D., Hansen, L.A., & Thal, L.J. (1998). Cognitive profiles of autopsy-confirmed Lewy body variant vs. pure Alzheimer disease. *Archives of Neurology*, *55*, 994–1000.

Craik, F.I., Byrd, M., & Swanson, J.M. (1987). Patterns of memory loss in three elderly samples. *Psychological Aging*, *2*, 79–86.

Daly, E., Zaitchik, D., Copeland, M., Schmahmann, J., Gunther, J., & Albert, M. (2000). Predicting conversion to Alzheimer disease using standardized clinical information. *Archives of Neurology*, *57*, 675–680.

DeBettignies, B.H., Mahurin, R.K., & Pirozzolo, F.J. (1993). Functional status in Alzheimer's Disease and multi-infarct dementia: A comparison of patient performance and caregiver report. *Clinical Gerontology*,*12*, 31–49.

Doody, R.S., Stevens, J.C., Beck, C., Dubinsky, R.M., Kaye, J.A., Gwyther, L., Mohs, R.C., & Thal, L.J. (2001). Practice parameter: Management of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, *56*, 1154–1166.

Elias, M.F., Beiser, A., Wolf, P.A., Au, R., White, R.F., & D'Agostino, R.B. (2000). The preclinical phase of alzheimer disease: A 22-year prospective study of the Framingham Cohort. *Archives of Neurology*, *57*, 808–813.

Flicker, C., Ferris, S.H., & Reisberg, B. (1991). Mild cognitive impairment in the elderly: Predictors of dementia. *Neurology*, *41*, 1006–1009.

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

Fowler, K.S., Saling, M.M., Conway, E.L., Semple, J.M., & Louis, W.J. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society*, *8*, 58–71.

Gainotti, G., Marra, C., Villa, G., Parlato, V., & Chiarotti, F. (1998). Sensitivity and specificity of some neuropsychological markers of Alzheimer dementia. *Alzheimer Disease Association Disorder*, *12*, 152–162.

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*, *51*, 888–895.

Goodglass, H. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia: Lea & Febiger.

Grundman, M., Corey-Bloom, J., Jernigan, T., Archibald, S., & Thal, L.J. (1996). Low body weight in Alzheimer's disease is associated with mesial temporal cortex atrophy. *Neurology*, *46*, 1585–1591.

Haxby, J.V., Grady, C.L., Koss, E., Horwitz, B., Schapiro, M., Friedland, R.P., Rapoport, S.I. (1988). Heterogeneous anterior-posterior metabolic patterns in dementia of the Alzheimer type. *Neurology*, *38*, 1853–1863.

Howieson, D.B., Dame, A., Camicioli, R., Sexton, G., Payami, H., & Kaye, J.A. (1997). Cognitive markers preceding Alzheimer's dementia in the healthy oldest old. *Journal of the American Geriatric Society*, *45*, 584–589.

Huff, F.J., Becker, J.T., Belle, S.H., Nebes, R.D., Holland, A.L., & Boller, F. (1987). Cognitive deficits and clinical diagnosis of Alzheimer's disease. *Neurology*, *37*, 1119–1124.

Hultsch, D.F., Hertzog, C., Small, B.J., McDonald-Miszczak, L., & Dixon, R.A. (1992). Short-term longitudinal change in cognitive performance in later life. *Psychological Aging*, *7*, 571–584.

Jack, C.R., Jr., Dickson, D.W., Parisi, J.E., Xu, Y.C., Cha, R.H., O'Brien, P.C., Edland, S.D., Smith, G.E., Boeve, B.F., Tangalos, E.G., Kokmen, E., & Petersen, R.C. (2002). Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology*, *58*, 750–757.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia: Lea & Febiger.

Killiany, R.J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., Tanzi, R., Jones, K., Hyman, B.T., & Albert, M.S. (2000). Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Annals of Neurology*, *47*, 430–439.

Lafleche, G. & Albert, M. (1995). Executive function deficits in mild Alzheimer's disease. *Neuropsychology*, *9*, 313–320.

Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychological Aging*, *8*, 207–220.

Masur, D.M., Fuld, P.A., Blau, A.D., Crystal, H., & Aronson, M.K. (1990). Predicting development of dementia in the elderly with the Selective Reminding Test. *Journal of Clinical and Experimental Neuropsychology*, *12*, 529–538.

Mattis S. (Ed.). (1976). *Mental status examination for organic mental syndrome in the elderly patient*. New York: Grune & Stratton.

Mayeux, R., Stern, Y., Rosen, J., & Leventhal, J. (1981). Depression, intellectual impairment, and Parkinson disease. *Neurology* , *31*, 645–650.

Mayeux, R., Stern, Y., & Spanton, S. (1985). Heterogeneity in dementia of the Alzheimer type: Evidence of subgroups. *Neurology*, *35*, 453–461.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*, 939–944.

Monsch, A.U., Bondi, M.W., Butters, N., Salmon, D.P., Katzman, R., & Thal, L.J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49*, 1253–1258.

Morris, J.C., McKeel, D.W., Jr., Storandt, M., Rubin, E.H., Price, J.L., Grant, E.A., Ball, M.J., & Berg, L. (1991). Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology*, *41*, 469–478.

Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*, 303–308.

Petersen, R.C., Stevens, J.C., Ganguli, M., Tangalos, E.G., Cummings, J.L., & DeKosky, S.T. (2001). Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). *Neurology*, *56*, 1133–1142.

Rabinowitz, J.C. & Craik, F.I. (1986). Prior retrieval effects in young and old adults. *Journal of Gerontology*, *41*, 368–375.

Ritchie, K., Touchon, J., Ledesert, B., Leibovici, D., & Gorce, A.M. (1997). Establishing the limits and characteristics of normal age-related cognitive decline. *Review de Epidemiologie et de Santé Publique*, *45*, 373–381.

Rosen, W. (1981). *The Rosen Drawing Test*. Bronx, NY: Veterans Administration Medical Center.

Rubin, E.H., Storandt, M., Miller, J.P., Kinscherf, D.A., Grant, E.A., Morris, J.C., & Berg, L. (1998). A prospective study of cognitive function and onset of dementia in cognitively healthy elders. *Archives of Neurology*, *55*, 395–401.

Salthouse T. (1989). *Mechanisms of age-cognition relations in adulthood*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schaie, K.W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychological Aging*, *4*, 443–453.

Scherr, P.A., Albert, M.S., Funkenstein, H.H., Cook, N.R., Hennekens, C.H., Branch, L.G., et al. (1988). Correlates of cognitive function in an elderly community population. *American Journal of Epidemiology*, *128*, 1084–1101.

Schwab, J. & England, A. (Eds.). (1969). *Projection technique for evaluating surgery in Parkinson's disease*. Edinburgh: E&S Livingstone.

Small, S.A., Stern, Y., Tang, M., & Mayeux, R. (1999). Selective decline in memory function among healthy elderly. *Neurology*, *52*, 1392–1396.

Stern, Y., Andrews, H., Pittman, J., Sano, M., Tatemichi T., Lantigua, R., & Mayeux, R. (1992). Diagnosis of dementia in a heterogeneous population. Development of a neuropsychological paradigm-based diagnosis of dementia and quantified correction for the effects of education. *Archives of Neurology*, *49*, 453–460.

Storandt, M. & Hill, R.D. (1989). Very mild senile dementia of the Alzheimer type. II. Psychometric test performance. *Archives of Neurology*, *46*, 383–386.

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*, *46*, 661–665.

Visser, P.J., Verhey, F.R.J., Hofman, P.A.M., Scheltens, P., & Jolles, J. (1999). Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. *Journal of Neurology*, *246*, 477–485.

Wallace, W.P. & Collins, J.E. (1991). The influence on recognition of spoken words that are misperceived. *Memory and Cognition*, *19*, 498–506.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised*. New York: The Psychological Corporation.

Wechsler D. (1987). *Wechsler Memory Scale–Revised*. New York: The Psychological Corporation.

Wolf, H., Grunwald, M., Kruggel, F., Riedel-Heller, S.G., Angerhfer, S., Hojjatoleslami, A., Hansel, A., Arendt, T., & Gertz, H. (2001). Hippocampal volume discriminates between normal cognition, questionable and mild dementia in the elderly. *Neurobiology of Aging*, *22*, 177–186.

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*, *48*, 278–281.

Appendix A

Barona's estimated verbal IQ = (54.23) + (0.49 × age category) + (1.92 × gender category) + (4.24 × race category) + (5.25 × education category) + (1.89 × occupation category) + (1.24 × urban/rural residence).