Published in final edited form as:

Clin Neuropsychol. 2012 ; 26(1): 18–30. doi:10.1080/13854046.2011.639310.

Normative Data in Women Age 85 and Older: Verbal Fluency, Digit Span, and the CVLT-II Short Form

Eric M. Fine¹, Joel H. Kramer¹, Li-Yung Lui², and Kristine Yaffe^{1,3,4} for the Study of Osteoporotic Fractures (SOF) Research Group

¹Department of Neurology, University of California, San Francisco, CA

²Research Institute, California Pacific Medical Center, San Francisco, CA

³Department of Psychiatry, University of California, San Francisco, CA

⁴Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

Abstract

Individuals age 85 years and above (i.e., the oldest old) represent the fastest growing segment of the U.S. population and are at increased risk of developing dementia. This represents an important challenge for the clinical neuropsychologist, as the extant normative data on neuropsychological measures remains relatively limited for this age group. Therefore, the aim of the present study was to characterize the performance effects of age and education in a large, well-characterized sample of women between the ages of 85 and 95 years on the CVLT-II Short Form (Delis et al., 2000), verbal fluency tasks, and the WAIS-III Digit Span Test (Wechsler, 1997). In order to minimize the likelihood that women with an incipient neurodegenerative process were included in the final normative sample, we applied regression-based change scores to identify and exclude women who evidenced a statistically significant decline on a global cognitive screening measure over a 20 year interval. The results of our analysis indicate varying influence of age and education on these measures and we provide tables with descriptive statistics stratified by both age and education. Findings from the present normative study are discussed within the context of “robust” longitudinal normative data.

Keywords

oldest old; normative data; CVLT

Introduction

Individuals aged 85 years and above, commonly referred to as the oldest old, are the fastest growing segment of the United States population, especially for women. The number of people in this age group is expected to increase by 40% in the US over the next decade alone (Moore, 2007). These changing demographics have significant public health ramifications due to the increasing proportion of healthcare needs and expenditures subsumed by this age range. The demographic shift, coupled with the evidence suggesting that the incidence of dementia doubles every five years (Jorm & Jolley, 1998), further constitutes a challenge for clinical neuropsychologists, as the corpus of normative data for many critical neuropsychological instruments remains insufficient in the oldest old (for review, see

Mitrushina, Boone, & D'Elia, 1999; Strauss, Sherman, & Spreen, 2006). The clinical implications of this are amplified in light of the fact that neuropsychologists are increasingly tasked with differentiating the subtle cognitive changes associated with an incipient neurodegenerative process (e.g., Mild Cognitive Impairment) from the effects of normal aging.

A significant challenge in developing normative data for older adults that accurately represent the effects of normal aging is the fact that individuals at a “preclinical” stage of dementia may score within normal limits on neuropsychological measures (Marcopulus & McLain, 2003; Saxton et al., 2004). For this reason, cross-sectionally derived normative data are inherently limited. To address this constraint, several recent studies have published “robust” longitudinal normative data on several neuropsychological measures (De Santi et al., 2008; Ritchie, Frerichs, & Tuokko, H, 2007; Sliwinski, Lipton, Buschke, & Steward, 1996; Pedraza et al., 2010; Holtzer et al., 2008). In this approach, longitudinal data is utilized for the purpose of excluding from the final normative sample individuals who subsequent to the baseline assessment transition into mild cognitive impairment (MCI) or dementia. These studies have tended to show “robust” normative samples yield higher test mean scores and reduce test variance when compared with conventional cross-sectional normative samples and consequently improve identification of individuals who ultimately transition to dementia (De Santi et al., 2008; Sliwinski et al., 1997; Holtzer et al., 2008; Pedraza et al., 2010).

Given that that the oldest old are at highest risk of developing dementia (Jorm & Jolley, 1998; Yaffe et al., 2011), the use of purely cross-sectional techniques for deriving normative data in this age cohort may be particularly problematic. Consider as an example a normative study by Whittle et al. (2007) of non-demented individuals age 90 and above. Although this study possesses a number of strengths (e.g., a large sample with normative reference scores provided for a range of neurocognitive measures), 53% of individuals included in the final normative sample were described as being “Cognitively-Impaired But Not Demented (CIND)”. As this study was purely cross-sectional, the normative sample may have included individuals an early stage of a neurodegenerative process. Thus these norms may be overstating the effects of normal aging on cognition, thereby diminishing their value for detecting a budding dementia condition.

The purpose of the present study, therefore, is to improve the extant normative data in the oldest old on a subset of commonly used neuropsychological measures. More specifically, we examined the effects of age and education on the California Verbal Learning Test-II Short Form (CVLT-II SF; Delis, Kramer, Kaplan, & Ober, 2000), measures of verbal generative fluency, and the digit span task (Wechsler, 1997) in a large sample of neurologically and cognitively intact women between the ages of 85 to 95 who were enrolled in a prospective, multi-site longitudinal study focused on health outcomes in women (Cummings et al., 1995). Bearing in mind the aforementioned limitations with conventional cross-sectional normative data, we incorporated longitudinal information into the sample selection criteria; this was accomplished by applying statistical criteria (i.e., standardized regression-based change scores) to exclude women who evidenced a significant performance decline over a twenty year period on a global measure of cognition (i.e., modified Mini-Mental Status Examination; mMMSE; Folstein, Folstein, & McHugh, 1975).

Methods

SOF/WISE Study

We studied women from the ongoing Study of Osteoporotic Fractures (SOF), a multi-center, prospective, observational study of women who were 65 years and older at baseline. There were 9,704 primarily white women recruited to the study between September 1986 and October 1988 from four metropolitan areas in the United States: Baltimore, Maryland; Minneapolis, Minnesota; Portland, Oregon; and Monongahela Valley, Pennsylvania. For this study, we included only the participants who attended the year 20 visit from 3 of the 4 clinic sites, completed the expanded cognitive battery, and had clinical cognitive status determined (Cummings et al., 1995). This ancillary study was called the Women, Cognitive Impairment Study of Exceptional Aging (WISE). All women provided written informed consent and the study was approved by the committees on human research at each study site and at the University of California, San Francisco.

Information regarding the participants' age and education was collected at baseline. At the year 20 visit, participants reported whether a doctor had ever diagnosed them with a variety of medical conditions. Further information on functional status was obtained at year 20 from caregivers or proxies using the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE; Jorm & Jacombe, 1989) and by a scale-based assessment of self-reported ability to perform activities of daily living (ADLs) and instrumental ADLs with or without difficulty (Yaffe et al., 2011). Briefly, for the IQCODE an informant rates the degree of change experienced by the examinee over the past 10 years with respect to everyday situations (e.g., "handling money for shopping"), with higher scores indicating greater decline. Depressive symptoms were evaluated at year 20 using the 15-point Geriatric Depression Scale (GDS). A score of 6 or higher was used to identify participants with depression (Lyneess et al., 1997).

Assessment of Cognitive Function

All SOF participants completed a 26-point shortened Mini-Mental State Examination (mMMSE; Folstein et al., 1975) every 2 to 4 years over the 20 year follow-up period. At Year 20, clinic staff administered a brief battery of neuropsychological measures; a more comprehensive evaluation was not feasible due to time constraints associated with the larger prospective study. The neurocognitive measures administered at Year 20, included the following: California Verbal Learning Test-II Short Form (CVLT-II SF; Delis et al., 2000), letter fluency task (words that start with the letter "F" over a one minute interval), category fluency task (vegetables over a one minute interval), and the Digit Span Test (Wechsler, 1997). Participants also completed at Year 20 the 3MS, (Teng & Chui, 1987), a modified version of the MMSE designed to sample a broader range of cognitive functions. The CVLT-II SF is nine-item version of the CVLT-II (Delis et al., 2000), in which a list of words are presented over four learning trials, followed by 30-second delayed free recall, 10-minute delayed free recall, 10-minute delayed cued recall, and 10-minute delayed recognition. At variance with the standardized test administration procedures, long-delay cued recall was not administered as part of the present study. This was done to reduce the time required to complete the cognitive assessment. Similarly, letter fluency was assessed using only a single letter rather than the more traditional three letter version in order to reduce test administration time. Although the use of a single letter fluency trial may diminish the reliability of the measure, previous studies have demonstrated that abbreviated fluency measures have utility for detecting early dementia and differentiating among different dementia etiologies (Canning, Leach, Stuss, Ngo, & Black, 2004; Kramer et al., 2003). For the purposes of the present study, normative references values were derived for

Letter and Category fluency, Digit Span Test (Wechsler, 1997) and the CVLT-II SF (Delis et al., 2000).

Screening for Cognitive Impairment

Participants were first screened for cognitive impairment. Those who met one or more of the following year 20 criteria were considered to screen positive for cognitive impairment: 1) Score of less than 88 on the 3MS (Teng & Chui, 1987), which is a suggested cutoff for screening for cognitive impairment in elderly using this measure (Espeland et al., 2006); 2) a score of 3.6 or greater on the IQCODE, the suggested cutoff for dementia screening in the elderly (Jorm & Jacombe, 1989); 3) a dementia diagnosis; or 4) living in a nursing home. Women who screened positive for cognitive impairment were clinically adjudicated for cognitive status (for additional details regarding the study methods, including cognitive status evaluation, please see Yaffe et al., 2011).

Clinical Cognitive Status Evaluation

A diagnosis was made by a randomly selected member of a panel of clinical experts, which included a neurologist, two neuropsychologists, and a geropsychologist. Information considered for the adjudication included year 20 neuropsychological scores medical history, medications, depression score, and current functional status.

The diagnosis of dementia was made based on DSM-IV criteria—that is, the development of multiple cognitive deficits that include memory impairment and at least one other cognitive disturbance (American Psychiatric Association, 1994). The cognitive deficit must be sufficiently severe to cause impairment in functioning and must be a decline from a previous level of functioning. MCI was diagnosed based on the presence of cognitive impairment (i.e., ~1.5 standard deviations below the mean for their age on a neuropsychological measure) that is insufficient to be dementia and along with relatively intact functional capacity.

Selection of the Final Normative Sample

Of the 1253 women from the original SOF-cohort who were between the ages of 85 and 95 at WISE, 261 were diagnosed with MCI and 217 were diagnosed with dementia. These participants were excluded from the normative sample. In addition, given the evidence that late-life depression is a risk factor for dementia (for review, see Byers & Yaffe, 2011), we also excluded from the analyses participants who obtained a score on the 15-item GDS of 6 or higher; this is the suggested cutoff for screening for depression in the elderly based on a study by Lyness et al (1997). In the remaining sample ($n = 726$), we adopted techniques described by McSweeney and colleagues for determining whether a patient has changed meaningfully over time on a cognitive measure (McSweeney, Naugle, Chelune, & Luders, 1993). Specifically, a standardized regression-based (SRB) equation was calculated with the model including as predictors age (years), education (years), and baseline mMMSE (Folstein et al., 1975), with the year 20 mMMSE score as the dependent measure. Based on the SRB equation, a predicted mMMSE score was calculated. We then transformed the difference between the observed mMMSE score at Year 20 and the predicted score into a standardized z -score using the following equation: $z\text{-score} = (Y_o - Y_p)/SE_{est}$, where Y_o is the observed Year 20 score, Y_p is the predicted score, and SE_{est} is the standard error of the estimate from the regression analysis. The SRB z -scores allows for an individual determination of change based on prevalence rates in the distribution. Consistent with studies in clinical populations (e.g., see McSweeney et al., 1993; Heaton et al., 2001), an SRB z -score of ± 1.64 (90% confidence band) was adopted to delimit a statistically significant change on the mMMSE from year 1 to year 20. After applying the SRB z -scores, we excluded an additional 46 participants.

Data Analysis

In the final normative sample, the impact of age and education on the following CVLT-II SF variables was assessed separately in regression analyses: Total words recalled on Trial 1, Trial 4, short-delay free recall, long-delay free recall, recognition hits, recognition false positives, and total recognition discriminability (calculated using equations provided in the CVLT-II manual; Delis et al., 2000). Similarly, the effects of age and education on digit span (forward, backward, and total digit span evaluated individually), letter fluency, and category fluency were evaluated using linear regression. In addition for each measure and/or condition, the sample was stratified according to age (85-86, 87-89, 90-95) and education (< 12 years of education and > 12 years of education); education was stratified in this manner to maintain cell sizes of at least 50, the suggested lower limit for cell size in normative data samples (Mitrushina et al., 1999). The main effect of age and education band for each cognitive measure was evaluated using one-way analysis of variance (ANOVA).

Results

The final sample was comprised of 680 women (see table 1 for demographic characteristics of final normative sample; table 2 includes information regarding medical co-morbidities), with 98% of the participants being of non-Hispanic, Caucasian origin. Regression analyses for the CVLT-II SF yielded a significant effect of age on Trial 1, Trial 4, Trial 1-4 total, short-delay free recall, long-delay free recall, and false positive errors on long-delay recognition. The effect of education on the CVLT-II SF was significant for the total number of words recalled on Trial 1, Trial 4, and Trial 1-4 total. Age significantly predicted performance on category fluency and there was a significant effect of education for letter fluency, category fluency, digit span forward total, and digit span backward total (see table 5 for results from regression analyses).

A similar set of findings emerged when evaluating group differences (based on age and education stratifications described above) using one-way analysis of variance (ANOVA). Specifically, significant main effects of age were obtained for the following CVLT-II SF conditions: trial 1 words recalled ($p = .003$), trial 4 words recalled ($p = .036$), total words recalled trials 1-4 ($p = .001$), and total false positive errors ($p = .047$). The main effect of age group for long-delay free recall trended towards significance ($p = .069$). The main effect of education was significant for the following CVLT-II SF conditions: trial 1 words recalled ($p = .0089$), trial 4 words recalled ($p = .009$), total words recalled trials 1-4 ($p = .001$), and short-delay free recall ($p = .042$). In terms of the other cognitive measures, significant main effects of age were obtained on category fluency ($p = .003$), with significant main effects of education observed on letter fluency ($p < .001$), category fluency ($p = .029$), digit span backward ($p = .009$), and digit span total ($p = .009$).

Discussion

Given that individuals age 85 years and above (i.e., the oldest old) represent the fastest growing segment of the US population and are at increased risk of developing dementia, there is burgeoning need for adequate normative data in this age cohort. In the present study, therefore, we examined the performance effects of age and education on the 9-item version of the CVLT-II (CVLT-II SF; Delis et al., 2000), verbal fluency measures, and the WAIS-III Digit Span Test (Wechsler, 1997) in neurologically intact and cognitively stable women between the ages of 85-95 years old. The statistical analyses revealed small (Cohen, 1988) but significant effects of age and/or education on all measures. In addition, the results further indicate that for certain variables (i.e., letter fluency and digit span) stratifying by education and not age is sufficient for normative data purposes in the oldest old. Having said

that, in the present study normative reference values for all neurocognitive variables are stratified by both age and education to simplify task (or condition) comparisons.

In comparison with the normative data on the CVLT-II SF from the Whittle et al. (2007) oldest old study and from the original standardization sample for this measure (Delis et al., 2000), the normative sample from the present study tended to obtain higher scores and was less variable. For example, based on the Delis et al. (2000) normative sample, a female between the ages of 80-89 who obtained a raw score of 18 on trials 1-4, would fall -1.5 standard deviation below average, a commonly used cut-off for impairment; in contrast, a score of 20 falls approximately at this level when applying norms from the present study. It should be noted that for letter fluency ("F" words) and digit span, our normative sample did not appear to outperform the Whittle et al. (2007) sample.

As noted in the introduction, conventional cross-sectional norms, such as those from Whittle et al. (2007) and from the original CVLT-II standardization study (Delis et al., 2000), are inherently constrained due to fact that individuals with an early neurodegenerative process may appear cognitive intact (i.e., still perform within normal limits on neuropsychological measures) and thus included in normative samples (for discussion, see Sliwinski et al., 1996). To help rectify this, several groups have recently published "robust" longitudinal normative data that exclude individuals from the final normative sample individuals that subsequent to the baseline assessment transition into MCI or dementia. These studies, collectively, have demonstrated that "robust" longitudinal norms tend to yield higher mean scores and lower score variability than those derived using conventional cross-sectional approaches; the end result being improved sensitivity to the early stages of a neurodegenerative process (De Santi et al., 2008; Sliwinski et al., 1996; Pedraza et al., 2010; Holtzer et al., 2008). Although for the present study we were not able to exclude individuals who may have transitioned into MCI or dementia subsequent when the neurocognitive measures were administered, we did, however, utilize longitudinal information in the form of regression-based change scores to confirm that the participants comprising the final normative sample were cognitively stable over a 20 year period. In addition, our final normative sample contained only those women who attended the final study visit (at year 20). This is important in light of the evidence from a recent "robust" longitudinal normative study that the inclusion of individuals who are "lost-to-follow-up" in a normative sample drives down test score means and increases test variability (Holtzer et al., 2008). These longitudinal features, therefore, may have reduced the number of preclinical dementia cases included in our final normative sample; the result being a more accurate representation of normal, non-pathological age-related changes on neuropsychological measures in the oldest old.

In addition, to the potential advantages conferred by the longitudinal features of our normative study, further merits of the present study include: (1) a large, well-characterized sample; and (2) performance on the neuropsychological measures was stratified based on both age and education, which is optimal for normative data. In terms of the study's limitation, given that we do not have longitudinal data to determine cognitive outcomes, it is unclear whether the normative reference scores from the present study improve the sensitivity of the cognitive measures to early dementia. Moreover, the nature of the sample selection, including the fact that the final normative sample only included individuals who maintained study participation across a 20 year period, may have biased the sample to a particular subset of elderly women (i.e., a "super" geriatric sample). As a result, the normative reference values obtained in the present study may not fully represent test score variability associated with normal aging. It is also important to recognize that the present normative study would clearly have been strengthened had we utilized a measure of learning/memory or executive function to quantify cognitive stability in the sample, rather

than a brief, and less sensitive, global cognitive screening measure. In addition, the present study is further limited by virtue of only including women and the fact the ethnic/racial composition of the sample was quite constrained (i.e., almost exclusively non-Hispanic, Caucasian women). Finally, the norms for long-delay recognition on the CVLT-II SF from the present study should clearly only be utilized if long-delay cued recall was not administered, as was done in our study.

In summary, the present study provides valuable normative data on commonly used neuropsychological measures in the oldest old and thus may enhance the identification of individuals in this age cohort who are exhibiting subtle cognitive deficits associated with an early neurodegenerative process.

Acknowledgments

The Study of Osteoporotic Fractures (SOF) and SOF-WISE is supported by the National Institutes of Health funding (AG05407, AR35582, AG05394, AR35584, AR35583, R01 AG005407, R01 AG027576-22, 2 R01 AG005394-22A1, and 2 R01 AG027574-22A1, 5R01AG026720-04). Dr. Kristine Yaffe is supported in part by the National Institute of Aging grant K24 AG 031155 and an Independent Investigator Award from the Alzheimer's Association. Dr. Kramer is a co-author of the CVLT-II and receives royalties from the test.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. American Psychiatric Association; Washington, DC: 2000.
- Byers AL, Yaffe K. Depression and risk of developing dementia. *Nature Reviews Neurology*. 2011; 7(6):323–331.
- Canning SJ, Leach L, Stuss D, Ngo L, Black SE. Diagnostic utility of abbreviated fluency measures in Alzheimer's disease and vascular dementia. *Neurology*. 2004; 62(4):556–561. [PubMed: 14981170]
- Cohen, J. Statistical power for the behavioral sciences. 2nd ed. Erlbaum; Hillsdale, NJ: 1988.
- Cummings SR, Nevitt MC, Browner WS, et al. Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. *New England Journal of Medicine*. 1995; 332:767–773. [PubMed: 7862179]
- Delis, DC.; Kramer, JH.; Kaplan, E.; Ober, BA. California Verbal Learning Test: Adult version (CVLT-II): Manual. 2nd ed. Psychological Corporation; San Antonio, TX: 2000.
- De Santi S, Pirraglia E, Barr W, Babb J, Williams S, Rogers K, et al. Robust and conventional neuropsychological norms: diagnosis and prediction of age-related cognitive decline. *Neuropsychology*. 2008; 22(4):469–484. [PubMed: 18590359]
- Espeland MA, Rapp SR, Robertson J, Granek I, Albert M, Bassford T, et al. Benchmarks for designing two-stage studies using modified mini-mental state examinations: experience from the Women's Health Initiative Memory Study. *Clinical Trials*. 2006; 3(2):99–106. [PubMed: 16773952]
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Heaton RK, Temkin N, Dikmen S, Avitable N, Taylor MJ, Marcotte TD, et al. Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology*. 2001; 16(1):75–91. [PubMed: 14590193]
- Holtzer R, Goldin Y, Zimmerman M, Katz M, Buschke H, Lipton RB. Robust norms for selected neuropsychological tests in older adults. *Archives of Clinical Neuropsychology*. 2008; 23:531–541. [PubMed: 18572380]
- Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine*. 1989; 19:1015–1022. [PubMed: 2594878]
- Jorm AF, Jolley D. The incidence of dementia: a meta-analysis. *Neurology*. 1998; 51(3):728–733. [PubMed: 9748017]

- Kramer JH, Jurik J, Sha SJ, Rankin KP, Rosen HJ, Johnson JK, et al. Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer's disease. *Cognitive and Behavioral Neurology*. 2003; 16:211–218. [PubMed: 14665820]
- Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Depression Scale. *Archives of Internal Medicine*. 1997; 157:449–454. [PubMed: 9046897]
- Marcopulus BA, McLain CA. Are our norms “normal”? A 4-year follow-up study of a biracial sample of rural elders with low education. *The Clinical Neuropsychologist*. 2003; 17(1):19–33. [PubMed: 12854008]
- McSweeney AJ, Naugle RI, Chelune GJ, Luders H. “T scores for change”: An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist*. 1993; 7:300–312.
- Mitrushina, MN.; Boone, KL.; D'Elia, L. Handbook of normative data for neuropsychological assessment. Oxford University Press; New York: 1999.
- Moore A. Older people. We can work it out. *Health Service Journal*. 2007; 117:24–26. [PubMed: 17319526]
- Pedraza O, Lucas JA, Smith GE, Petersen RC, Graff-Radford NR, Ivnik RJ. Robust and expanded norms for the Dementia Rating Scale. *Archives of Clinical Neuropsychology*. 2010; 25(5):347–358. [PubMed: 20427376]
- Ritchie L, Frerichs R, Tuokko H. Effective normative samples for the detection of cognitive impairment in older adults. *The Clinical Neuropsychologist*. 2007; 21(6):863–874. [PubMed: 17853155]
- Saxton J, Lopez OL, Ratcliff G, Dulberg C, Fried LP, Carlson MC, et al. Preclinical Alzheimer disease: neuropsychological test performance 1.5 to 8 years prior to onset. *Neurology*. 2004; 63(12):2341–2347. [PubMed: 15623697]
- Sliwinski M, Lipton RB, Buschke H, Steward W. The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journal of Gerontology: Psychological Sciences*. 1996; 51:217–225.
- Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests. 3rd ed. Oxford University Press; New York: 2006.
- Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *Journal of Clinical Psychiatry*. 1987; 48:314–318. [PubMed: 3611032]
- Wechsler, D. Wechsler Adult Intelligence Scale-III. The Psychological Corporation; San Antonio, TX: 1997.
- Whittle C, Corrada MM, Dick M, Ziegler R, Kahle-Wroblewski K, Paganini-Hill A, et al. Neuropsychological data in nondemented oldest old: the 90+ Study. *Journal of Clinical and Experimental Neuropsychology*. 2007; 29(3):290–299. [PubMed: 17454349]
- Yaffe K, Middleton LE, Lui L, Spira AP, Stone K, Racine C, et al. Mild cognitive impairment, dementia, and their subtypes in oldest old women. *Archives of Neurology*. 2011; 68(5):631–636. [PubMed: 21555638]

Table 1

Demographics of normative sample and scores on the mMMSE, Teng 3MS, and GDS

Age range	N	%	mMMSE year 1 (mean/std)	mMMSE year 20 (mean/std)	Teng 3MS (mean/std)	15-item GDS (mean/std)
85-86	260	38.2	25.24 (1.08)	25.03 (1.09)	93.92 (3.35)	1.58 (1.41)
12 yrs ed.	143	45.0				
> 12 yrs ed.	117	55.0				
87-89	279	41.0	25.25 (1.23)	24.93 (1.11)	93.05 (3.56)	1.64 (1.39)
12 yrs ed.	148	53.0				
> 12 yrs ed.	131	47.0				
90-95	141	20.7	25.18 (1.23)	24.50 (1.34)	92.14 (3.55)	1.92 (1.50)
12 yrs ed.	89	63.8				
> 12 yrs ed.	52	36.2				

Abbreviations: modified Mini-Mental State Examination (mMMSE); Teng Modified Mini-Mental State Examination (3MS); Geriatric Depression Scale (GDS)

Table 2

Overall frequency of medical co-morbidities and self-reported depression

Medical History	No.	%
Heart Disease ^a	157	23.1
Stroke	74	11.0
Diabetes	85	12.5
Cancer	167	24.5

^aIncludes: Coronary artery disease, myocardial infarction, congestive heart failure

Table 3

Results from regression analyses (separate models for age and education)

Cognitive Measure	<u>Age</u>			<u>Education</u>		
	β	t	p	β	t	p
CVLT-II SF						
Trial 1	-.133	-3.50	<.001	.081	2.15	.035
Trial 4	-.110	-2.88	.004	.096	2.50	.012
Trial 1-4 total	-.144	-3.78	<.001	.121	3.16	.002
SDFR	-.088	-2.30	.022	.062	1.62	.105
LDFR	-.105	-2.76	.006	.054	1.41	.159
LDR Hits	.018	.477	.633	-.048	-1.25	.211
LDR FPs	.092	2.39	.017	.007	.172	.863
LDRD	-.047	-1.21	.224	-.027	-.695	.487
Letter Fluency	.001	-.012	.990	.161	4.22	<.001
Category Fluency	-.099	-2.58	.010	.058	1.51	.129
Digit Span Forward	-.014	-.376	.707	.071	1.84	.065
Digit Span Backward	-.062	-1.616	.106	.167	4.38	<.001
Digit Span Total	-.045	-1.18	.237	.139	3.64	<.001

Note: short-delay free recall (SDFR); long-delay free recall (LDFR); long-delay recognition (LDR); long-delay recognition discriminability (LDRD)

Table 4

Means and standard deviations for the CVLT-II SF stratified by age and education

CVLT-II SF Condition	Age/Education Category	Mean	SD
Trial 1 total ++	85-86 yrs	4.8	1.2
	12 yrs ed.	4.8	1.2
	> 12 yrs ed.	4.9	1.1
	87-89 yrs	4.6	1.4
	12 yrs ed.	4.4	1.4
	> 12 yrs ed.	4.8	1.4
	90-95 yrs	4.4	1.3
	12 yrs ed.	4.3	1.4
	> 12 yrs ed.	4.5	1.3
Trial 4 total ++	85-86 yrs	7.7	1.2
	12 yrs ed.	7.6	1.3
	> 12 yrs ed.	7.7	1.1
	87-89 yrs	7.5	1.2
	12 yrs ed.	7.3	1.2
	> 12 yrs ed.	7.7	1.0
	90-95 yrs	7.4	1.3
	12 yrs ed.	7.3	1.3
	> 12 yrs ed.	7.4	1.3
Trial 1-4 total ++	85-86 yrs	26.3	3.9
	12 yrs ed.	26.0	4.2
	> 12 yrs ed.	26.7	3.7
	87-89 yrs	25.5	4.3
	12 yrs ed.	24.8	4.4
	> 12 yrs ed.	26.3	4.0
	90-95 yrs	24.8	4.3
	12 yrs ed.	24.6	4.3
	> 12 yrs ed.	25.0	4.3
Short-Delay Free Recall +	85-86 yrs	7.1	1.4
	12 yrs ed.	6.9	1.4
	> 12 yrs ed.	7.2	1.3
	87-89 yrs	7.1	1.5
	12 yrs ed.	6.9	1.5
	> 12 yrs ed.	7.2	1.4
	90-95 yrs	6.8	1.4
	12 yrs ed.	6.8	1.5
	> 12 yrs ed.	6.7	1.4
	85-86 yrs	6.5	1.7

CVLT-II SF Condition	Age/Education Category	Mean	SD
Long-Delay Free Recall	12 yrs ed.	6.4	1.8
	> 12 yrs ed.	6.6	1.7
	87-89 yrs	6.4	1.8
	12 yrs ed.	6.2	1.8
	> 12 yrs ed.	6.6	1.7
	90-95 yrs	6.1	1.9
	12 yrs ed.	6.2	1.9
	> 12 yrs ed.	5.9	2.0
Recognition Hits	85-86 yrs	8.4	1.0
	12 yrs ed.	8.5	.77
	> 12 yrs ed.	8.2	1.2
	87-89 yrs	8.4	.90
	12 yrs ed.	8.4	.80
	> 12 yrs ed.	8.4	.90
	90-95 yrs	8.4	1.0
	12 yrs ed.	8.3	.95
Recognition FPs *	> 12 yrs ed.	8.4	1.3
	85-86 yrs	1.0	1.0
	12 yrs ed.	1.0	1.0
	> 12 yrs ed.	.85	.95
	87-89 yrs	1.0	1.3
	12 yrs ed.	1.0	1.1
	> 12 yrs ed.	1.1	1.4
	90-95 yrs	1.3	1.4
Long-Delay RD	12 yrs ed.	1.1	1.2
	> 12 yrs ed.	1.5	1.6
	85-86 yrs	2.9	.52
	12 yrs ed.	3.0	.48
	> 12 yrs ed.	2.9	.55
	87-89 yrs	2.9	.50
	12 yrs ed.	2.9	.49
	> 12 yrs ed.	2.9	.50
	90-95 yrs	2.7	.58
	12 yrs ed.	2.9	.53
	> 12 yrs ed.	2.8	.65

Note: Long-Delay Cued-Recall was not administered ; False positives (FPs); recognition discriminability (RD);

++ significant group differences for both age and education;

* significant group differences for age;

+ significant group differences for education

Table 5

Means and standard deviations for the verbal fluency task and the WAIS-III Digit Span Test stratified by age and education

Cognitive Measure	Age/Education Category	Mean	SD
Letter Fluency ("F" words) +	85-86 yrs	11.6	4.2
	12 yrs ed.	11.3	4.4
	> 12 yrs ed.	12.0	4.0
	87-89 yrs	11.7	3.9
	12 yrs ed.	11.2	3.8
	> 12 yrs ed.	12.5	3.8
	90-95 yrs	11.6	3.5
	12 yrs ed.	11.1	3.2
	> 12 yrs ed.	12.6	3.7
Category Fluency (vegetables) ++	85-86 yrs	12.0	3.0
	12 yrs ed.	11.7	2.8
	> 12 yrs ed.	12.4	3.1
	87-89 yrs	11.9	3.2
	12 yrs ed.	11.5	3.0
	> 12 yrs ed.	12.3	3.3
	90-95 yrs	11.0	2.8
	12 yrs ed.	10.6	2.7
	> 12 yrs ed.	11.2	2.7
Digit Span Forward Total	85-86 yrs	7.6	2.1
	12 yrs ed.	7.6	2.0
	> 12 yrs ed.	7.6	2.2
	87-89 yrs	7.4	1.9
	12 yrs ed.	7.3	2.0
	> 12 yrs ed.	7.5	1.9
	90-95 yrs	7.5	2.2
	12 yrs ed.	7.4	2.2
	> 12 yrs ed.	7.8	2.3
Digit Span Backward Total +	85-86 yrs	6.1	2.0
	12 yrs ed.	6.1	1.9
	> 12 yrs ed.	6.1	2.1
	87-89 yrs	6.0	2.0
	12 yrs ed.	5.8	2.0
	> 12 yrs ed.	6.4	2.0
	90-95 yrs	5.9	1.8
	12 yrs ed.	5.6	1.7
	> 12 yrs ed.	6.2	2.0

Cognitive Measure	Age/Education Category	Mean	SD
Digit Span Total +	85-86 yrs	13.8	3.5
	12 yrs ed.	13.8	3.5
	> 12 yrs ed.	13.8	2.6
	87-89 yrs	13.5	3.3
	12 yrs ed.	13.1	3.2
	> 12 yrs ed.	13.9	3.3
	90-95 yrs	13.5	3.6
	12 yrs ed.	13.0	3.4
	> 12 yrs ed.	14.0	3.8

⁺⁺ *Symbols:* significant group differences for both age and education;

⁺ significant group differences for education