Published in final edited form as:
Clin Neuropsychol. 2012 May ; 26(4): 641–653. doi:10.1080/13854046.2012.666266.

Normative Data for Neuropsychological Tests in a Rural Elderly Chinese Cohort

Lili Yang[1], **Frederick W. Unverzagt, Ph.D.**[2], **Yinlong Jin, M.D.**[3], **Hugh C. Hendrie, M.D., D.Sc.**[2,7,8], **Chaoke Liang, B.S.**[3], **Kathleen S. Hall, Ph.D.**[2], **Jingxiang Cao, B.S.**[3], **Feng Ma, B.S.**[3], **Jill R. Murrell, Ph.D.**[4], **Yibin Cheng, B.S.**[3], **Ping Li, B.S.**[5], **Jianchao Bian, B.S.**[6], and **Sujuan Gao, Ph.D.**[1]

[1]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, Indiana

[2]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana

[3]Institute for Environmental Health and Related Product Safety, Chinese Center for Disease Control and Prevention, Beijing, China

[4]Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, Indiana

[5]Sichuan Provincial Center for Disease Control and Prevention in China, Chengdu, China

[6]Shandong Institute for Prevention and Treatment of Endemic Disease in China, Jinan, China

[7]Indiana University Center for Aging Research, Indianapolis, Indiana

[8]Regenstrief Institute, Inc. Indianapolis, Indiana

Abstract

Normative information is important for appropriate interpretation of cognitive test scores as a critical component of dementia diagnosis in the elderly population. A cross-sectional evaluation of 1826 participants aged 65 years and older from four rural counties in China was conducted using six cognitive instruments including tests of global cognitive function (the Community Screening Instrument for Dementia), Memory (Word List Learning and Recall tasks from the Consortium to Establish a Registry for Alzheimer's Disease, IU Story), Language (Animal Fluency Test), and executive function (IU Token). Multiple regression models adjusting for demographic variables were used to provide standardized residuals *z*-scores and corresponding percentile ranking for each cognitive test. In all cognitive tests, older age was associated with worse test performance while exposure to education was related to better cognitive test performance. We also detected a significant gender difference with men scoring better than women and a significant gender by education interaction on two tests. The interaction indicates that gender difference in test scores was much smaller in participants with more education than those who had less or no education. These demographically adjusted, regression-based norms can be a useful tool to clinicians involved with differential diagnosis of cognitive and memory disorders in older adults in rural China.

Keywords

Normative Study; Neuropsychological Test; Age; Gender; Education; Regression- Based Norms

---

Address correspondence to: Sujuan Gao, Ph.D., Department of Biostatistics, Indiana University School of Medicine, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202, Phone: (317) 274-0820, Fax: (317) 274-2678, sgao@iupui.edu.

Disclosure: The authors report no conflicts of interest.

Introduction

It is estimated that 24 million people worldwide have dementia with the number rising to 81 million by 2040 (Ferri, Prince, Brayne, Brodaty, Fratiglioni et al., 2005). The largest proportion of new dementia cases is expected to come from China and the developing western Pacific region where 1.2 million new cases per year are anticipated (Ferri et al., 2005). Cognitive assessment is routinely conducted as a critical component of dementia diagnosis. It is well known that differences in extent and approach to formal education, vocational experiences, interests, and engagement in reading and writing in the lifespan can combine in complex ways to produce locally distinctive influences on measured cognitive ability in older adults in any particular region (Hertzog, 1996; Anstey & Christensen, 2000). Hence, normative data that can take into account these demographic and background information is necessary for the interpretation of cognitive test scores in populations of interest.

Several normative studies in older adults of Asian descent have been published including those from Mainland China (Hsieh & Tori, 2007; Sosa, Albanese, Prince, Acosta, Ferri et al., 2009), Singapore (Lim, Collinson, Feng & Ng, 2010) and Taiwan (Lin, Wang, Liu, Chen, Lee et al., 2002). However, these studies either had small sample sizes or included only urban populations with the exception of one study reporting on neuropsychological performance in a relatively large sample of rural elderly participants (n=946) from a suburban county of Beijing, China (Sosa et al., 2009). More importantly, none of the studies examined potential interactions among demographic variables on cognitive performances. Given that the majority of the elderly Chinese population resides in rural areas, there is a need for neuropsychological data from large cohorts of community-based rural subjects. In this paper, we report on the joint effects of age, education, and gender on neuropsychological test performance in a sample of 1826 rural elderly Chinese men and women and provide demographically-adjusted, regression-based norms.

Method

Study Population

A total of 2000 older adults were recruited for the Selenium and Cognitive Decline study, a longitudinal epidemiologic project funded by the National Institute of Health examining the long-term impact of selenium on cognitive function in rural elderly Chinese [see Gao, Jin, Hall, Liang, Unverzagt et al. 2007 for details]. Two counties from Sichuan Province in southwestern China and two counties from Shandong Province in eastern China were selected based on: rural areas with historically low in- and out-migration, wide range of selenium levels but comparable trace elements otherwise, comparable age and gender proportions, and sufficient population to provide 500 elderly subjects per county. Sites with known endemic diseases (including Keshan disease, Kaschin-Beck disease, goiter and Cretinism, and fluorosis) were excluded from consideration.

Residents aged 65 years or older from these four counties were invited to participate in the study. The only exclusion criterion was examiner-rated hearing loss precluding cognitive assessment. Interviewers, who were employees of county-level Center for Disease Control offices, traveled to the sites, established a temporary headquarter, and conducted a complete door-to-door census of residents age 65 years and older in the area. Enrollment at each site continued from village to village until 500 elderly participants were enrolled. There were no refusals; all subjects provided written informed consent to participate in the study which was approved by both the Indiana University Purdue University Indianapolis Institutional

Review Board and the Institute for Environmental Health and Related Safety, Chinese Center for Disease Control and Prevention.

Cognitive Assessment

Cognitive tests used in this project were selected because of wide use in other large scale epidemiological studies of cognitive aging in US (Hall, Ogunniyi, Hendrie, Osuntokun et al., 1996; Ganguli, Belle, Ratcliff, Seaberg, Huff et al., 1993; Snitz, Unverzagt, Chang, Vander Bilt, Gao et al., 2009; Langa, Plassman, Wallace, Herzog, Heeringa et al., 2005); Africa (Hall et al., 1996); South America (Sosa et al., 2009), and Asia (Emsley, Gao, Li, Liang, Ji et al., 2000; Sosa et al., 2009). Cognitive assessment was conducted in face-to-face interviews in the homes of study participants.

Global Cognitive Status was measured with the Community Screening Instrument for Dementia (CSID) (Hall et al., 1996). The CSID has items assessing orientation, memory, abstract thinking, judgment, language, and praxis. For this project, a raw score method was used for scoring resulting in score range of 0–30 with higher scores indicating better cognitive function. The CSID has good two-week test-retest reliability (intraclass correlation=0.79) and inter-rater reliability (kappa=1 for 94% of the items) (Hall, Hendrie, Brittain, et al., 1993; Hall, Ogunniyi, Hendrie et al., 1994; Hall et al., 1996; Hall, Gao, Emsley, Ogunniyi, Morgan et al., 2000). Memory was assessed with two measures. Word List Learning (Morris, Heyman, Mohs, Hughes, van Belle et al., 1989) is a 3-trial, 10-item word list learning test with free recall taken after each learning trial and after a brief delay. Scores are the total number of words recalled across the three learning trials (range 0–30) and after delay (range 0–10) with higher scores indicating better performance. Word List Learning has a one-month test-retest reliability coefficient of 0.62 while Word List Delayed Recall has test-retest reliability of 0.64 (Morris et al., 1989). The IU Story is a prose recall task created for this project to be suitable to the Chinese culture and the rural population. The examiner reads a paragraph-length story out loud to the subject who then attempts to restate the story immediately. The story has 14 units of information that are gist scored (range 0–14) with higher scores indicating better memory. The story was tested in 1016 elderly Chinese in a previous pilot study and was found to be acceptable to the villagers and produced a normally distributed range of scores (Emsley et al., 2000). Language was assessed with the Animal Fluency Test (Isaacs & Akhtar, 1972) in which subjects name as many animals as possible in 60 seconds. One month test-retest was 0.67 in normal subjects without dementia (Morris et al., 1989). Executive Function was measured with the IU Token Test (Yamamoto, Evans, Johnson & Unverzagt, 2003; Snitz et al., 2009). A sheet of paper with an array of circles and squares of varying size (small and large) and color (red, black, yellow, and green) is presented to the subject. The examiner reads aloud a series of commands that ask the subject to point to the figures in various combinations. IU Token score is number correct (range 0–24) with higher scores indicating better performance.

The test items and instructions were harmonized to Chinese culture. The questionnaire was translated into Chinese, and back-translated into English. The translations were performed by persons outside the Beijing and Indiana University research teams in order to avoid bias.

Interviewers had college degrees in public health or related areas and spoke the local dialect. All interviews were conducted in the local dialect. Interviewers received training in all study instruments. Prior to conducting fieldwork, interviewers had to achieve greater than or equal to 95% inter-rater reliability in mock assessments done with volunteers from the community.

Other Information

Other information collected during the evaluation included age, gender, whether the participant attended school and years of schooling, marital status, history of cancer, diabetes, hypertension, stroke, heart attack, head injury and fracture by self-report.

Statistical Analysis

Group differences were evaluated using two sample *t*-test for continuous variable and Chi-square test for categorical variable. Q-Q plots for normality and residual plots for constant variance indicated that assumptions for multiple regression models were met. Following Parmenter et al. (Parmenter, Testa, Schretlen, Weinstock-Guttman & Benedict, 2010) and Smerbeck et al. (Smerbeck, Parrish, Yeh, Hoogs, Krupp et al., 2011), multiple regression models were used with each of the cognitive tests as the dependent variable, age, gender, education and all interactions as potential covariates. The final regression model consists of covariates and interactions with $p<0.05$. These models can then be used by readers to generate demographically-adjusted z-scores and percentiles to assist in interpreting the level of performance of any given patient as follows: 1) generate a demographically-adjusted predicted test score for a patient by entering the patient's demographic information into an equation using coefficient estimates from our multiple regression models, 2) subtract the obtained test score from the predicted test score and divide by the standard deviation of the residuals, 3) the resultant z-score can be interpreted directly or converted to a percentile which indicates the patient's standing relative to peers of comparable age, gender, and education. See supplementary data for spreadsheets that facilitate these calculations.

Results

Two thousand elderly subjects were recruited for the selenium project. For this normative study, we excluded participants with diseases that may interfere with cognitive performance. Hence 174 participants (18 with Parkinson's disease, 61 with stroke, and 95 with history of head injury) were excluded from data analysis. Demographic characteristics of the study participants are presented in Table 1. For the 1826 participants included in this analysis, mean age was 71.9 (SD = 5.6), mean years of education was 1.6 (SD = 2.6) with 62 % never attended school. Sixty-two percent were married, and hypertension was the most common self-reported medical disorder (15.7%). There were some notable differences by gender with women having significantly less education, higher rate of being widowed, and higher rate of diabetes.

Cumulative frequency distributions of raw scores for each cognitive test in the entire sample are presented in Table 2. These percentiles represent the performance ranking of test scores without taking into account any participant characteristics. We conducted a series of regression models using each cognitive score as the dependent variable, age, gender and education are considered as potential covariates as well as all two-way interactions. Regression coefficients, standard errors of the coefficients, p-values, standardized coefficient estimates, R square and standard deviations of residuals from the final regression models including only significant covariates and interactions are presented in Table 3. The three demographic variables (age, gender and education) accounted for the largest variation on the CSID (R^2=23%) and they accounted for the least variation on the IU story (R^2 =9%). All cognitive scores were affected by age (older age associated with worse cognitive performance), gender (men performed better than women), and education (more education associated with better cognitive performance). We also found significant interaction effects on the CSID and IU Token Test, indicating a greater gender difference in cognitive scores in those with less or no education than in those with more education. Main effects of age, education, and gender were observed for Word List Learning, Word List Recall, and Animal

Fluency. Main effects for age and education were noted for IU Story. We repeated the analysis using all 2000 subjects and the results did not differ much– i.e., parameter estimates unchanged at the 1st or 2nd decimal.

The figure displays an example of a spreadsheet setup that automatically calculates demographically-adjusted normative rankings (z-score and percentile) for the CSID indicating relative placement of an individual's observed score versus this sample. In the spreadsheet, column A lists variable names (e.g., cell A2 is the variable name, age). In column B, rows B2 to B6 are used to enter demographic information and observed CSID score from the individual (For example, cell B2 is for entering the actual value of age for the individual); rows B8 to B13 include parameter estimates and standard deviation of residuals from the multiple regression model of the CSID from Table 3 (e.g., the value in cell B9 is the coefficient estimate for age effect); the last three rows are used to generate predicted CSID score, z-score as well as the corresponding percentile ranking (the value in cell B14 is the predicted CSID score calculated using multiple regression model formula (1), cell B15 is the residual z-score generated by formula (2), and cell B16 is the corresponding percentile ranking using the statistic function NORMSDIST in Excel). As an illustration, we consider a female subject aged 71 years with 6 years of education and an observed CSID score of 26. Using the coefficients for CSID (row 3 to row7 in column 3 from Table 3), the multiple regression data can be combined to form an equation that generates a predicted CSID score for this individual as follows:

$$\text{CSID}_{\text{predicted}} = 38.5114 - 0.1738 \times \text{Age} - 1.9156 \times \text{Female} + 0.1811 \times \text{Years of Education} + 0.3203 \times \text{Female} \times \text{Years of Education}. \quad (1)$$

The demographically-adjusted predicted CSID score for this 71 year old, female, with 6 years of education is 27.26. By calculating the difference between observed (26) and predicted scores (27.26) divided by the standard deviation of residual (3.1198, value in row 7and column 8from Table 3), a demographically-adjusted residual *z*-score of −0.41 is generated using the following formula

$$\text{CSID z-score} = \frac{\text{CSID Score}_{\text{observed}} - \text{CSID Score}_{\text{Predicted}}}{\text{SD}_{\text{residuals}} \text{ of CSID}}. \quad (2)$$

This *z*-score can then be transformed into a percentile based on the normal distribution of the residuals from the regression model. For the subject in our illustration, a residual z-score of −0.41 equates to approximately the 34th percentile, a performance in the low end of the average range.

Discussion

We examined the effect of demographic variation on cognitive test performance in a large sample of rural Chinese older adults and also provided demographically adjusted regression based norms. In this large sample of older, rural Chinese adults, age and education had strong and consistent effects on cognitive test performance. For all six cognitive tests used in our study, older age was associated with worse test performance while education was positively related to cognitive test performance. Gender was also related to performance (men did better than women); however, this effect may be driven by a gender and education interaction in that the gender difference in test scores was much smaller in participants with more education than those who had less or no education. These demographically-adjusted regression-based norms will allow efficient ranking of an individual's cognitive scores based on this large sample thus facilitating test score interpretation.

Over the last two decades, many cognitive normative studies of older adults have been published. Busch and Chapin provided an excellent review of cognitive normative studies for older adults in North America and Europe (Busch & Chapin, 2008). Measures included in the review included the Mini-mental State Examination (MMSE), Modified Mini-mental State Examination (3MS), Dementia Rating Scale, Cognistat, CERAD cognitive battery, Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), and the Mayo Clinic normative studies projects. The review identified four large studies with sample sizes of 1000 or more. Ganguli et al. reported norms of the neuropsychological tests of the assessment protocol used in the Consortium to Establish a Registry for Alzheimer's Disease (CERAD), based on a rural community sample of 1350 subjects aged 65 years and older with median education of 12 years living in Pennsylvania (Ganguli, Ratcliff, Huff, Kancel, Fisher et al., 1991). They reported younger age, more education and female gender to be independently related to better performance on nearly all of the CERAD tests. Crum et al. reported age- and education- specific normative data for the Mini-Mental State Examination (MMSE) from a sample of 18056 adults aged 18 and older in US, with 46% participants receiving high school education and 32% participants having college or higher education (Crum, Anthony, Bassett & Folstein, 1991). They found older age and less education were associated with worse MMSE scores. Schmidt et al. reported norms for the Dementia Rating Scale (DRS) (Schmidt, Freidl, Fazekas, Reinhart, Grieshofer et al., 1994) based on a sample of 1001 healthy volunteers in Austria, aged 50 to 80 years, with mean of 10.8 years of schooling (68% had 10 or more years of education). They also observed participants with younger age and more education to have better performance on DRS test. In a study reported by Bravo and Hebert, age- and education-specific reference values for MMSE and 3MS Examination were derived from a Canadian population of 7754 subjects aged 65 years and older, 69% of whom had 9 or more years of schooling (Bravo & Hebert, 1997). In that study, MMSE and 3MS showed negative relationships with age and positive relationships with education. In addition, females scored slightly better than males. Among these four large studies, only the Ganguli et al., 1991 study examined performance of multiple neuropsychological tests, and none of them examined interactions among demographic variables except the Canadian study where age by education interactions were examined. In addition, none of these studies provided regression-based equations. Our results on the association between age and education and cognitive test performance are consistent with these four studies while our gender finding that men perform better than women was counter to the Ganguli et. al. report where women performed better than men. The reason for this reversal is unclear but it may be related to differences in the pervasiveness of education in the two settings, i.e., minimal in our China sample and extensive in the Ganguli sample. In situations where education is widely distributed in a sample, it may then be possible to see gender main effects that are not otherwise detectable and not confounded by education.

There are also several large studies reporting cognitive norms in elderly Asian populations. In a sample of 405 healthy Japanese adults aged 65 years and older with an average of 8 years of education, older age and less education were found to correspond to poorer performance on the Benton Visual Form Discrimination test (Kasai, Ishizaki, Ishii, Yamaguchi, Yamadori et al., 2009). Another normative study of 2266 elderly Japanese participants, 96% of whom had 6 or more years of schooling, showed the same trend of age negatively and education positively affecting MMSE scores (Ishizaki, Meguro, Ambo, Shimada, Yamaguchi, Hayasaka et al., 1998). Lin and colleagues examined the performance of the Cognitive Abilities Screening Instrument (CASI) in a sample of 1178 non-demented older Taiwanese adults, 57% of whom never attended to school (Lin et al., 2002). Again, an inverse relationship with age and a positive relationship with education were observed. In addition, Lin et al reported that among subjects with no education males scored significantly higher than females. The 10/66 Dementia Research Group included 946 rural elderly Chinese aged 65 and over, with 57% participants having no education, from a suburb in

Beijing and examined cognitive performances on a 3-trial 10-item word list, verbal fluency, and global cognitive screener (CSI-D) (Sosa et al., 2009). An urban sample of 1076 Chinese elders were also included in the 10/66 study. For both urban and rural samples, independent effects were noted for age, education, and gender. Increasing age and decreasing education were associated with lower scores and men tended to do better than women. Lee et. al studied the performance of the CERAD neuropsychological battery in 618 healthy Korean elderly aged 60 to 90 years where 26% of participants had no education. In their study, age was negatively related to performance on all tests and education was positively related to all tests except Word List Recognition. A gender by education interaction was found for Boston Naming Test, MMSE and Constructional Praxis such that the men had a performance advantage at lower education levels. Our results for the effects of age and education on cognitive test performance are consistent with these findings from Asia and specifically in line with those of Lee et al., 2004 in detecting a gender by education interaction where the male advantage is greatest at the lower levels of education.

This gender by education interaction suggests that the work and life roles of older Asian men and women with no education differed in a way that may have allowed men but not women to develop some cognitive or literacy skills. As children and adults, the men and women in this cohort would have experienced quite different paths. For the most part, women who did not attend school would have tended to household duties from a very young age to older adulthood, such as weaving, cooking, and other household chores. Men on the other hand would have greater opportunities working outside the home environment and interacting with others as the primary income earner for the family. It is possible that the life of men in this context is sufficiently interactive or demanding that they become self-educated, whereas the life of uneducated women is less diverse and stimulating with resultant differences in objective test performance. When education is more widely available, women receive the same beneficial exposure as men and do just as well cognitively if not better (see Ganguli et al., 1991).

The present study has a number of strengths including a large sample size from an understudied population and examination of a range of cognitive domains including memory, language, and executive functions. Our study population included participants with chronic diseases such as hypertension or diabetes, common comorbidities in the elderly population (Østbye, Krause, Norton, Tschanz, Sanders et al., 2006). It is important that normative samples reflect these medical complexities so that patients from similar populations can be appropriately evaluated. Our cognitive measures included several that were developed for populations with low literacy rate. By providing regression based norms on multiple cognitive tests from the same sample, we facilitate the comparisons between cognitive tests and avoid the inaccuracies resulting from norms derived from multiple samples.

Our study also has limitations. Relatively few men and almost no women from this cohort had more than 6 years of education. Despite this range restriction, we were able to detect an education effect and in fact to document an interesting gender-education interaction. We did not formally assess reading levels in this cohort. It is possible that some older adults with no formal schooling did learn to read and contrariwise that some who had exposure to formal schooling in their youth never applied reading in their daily lives. This would have the effect of attenuating the relationship of education to test performance in this study. Also, the cognitive battery we used did not include a measure of visuospatial function.

This study provides demographically-adjusted, regression-based norms that were derived from the largest sample of rural Chinese to date. Age and education had relatively strong influences on cognitive test performance while an interesting gender-education interaction suggested that gender differences were maximal among those with least amounts of education. The demographically-adjusted scores will be useful to clinicians involved in differential diagnosis of cognitive and memory disorders in older rural Chinese. These data will be most useful when the procedures we used in administration and scoring are followed and the patient in question matches closely with this sample demographically. To the degree there is significant divergence in assessment approach or demographic match, these reference values will have decreasing or even limited utility.

Acknowledgments

The research is support by NIH grant R01 AG019181, R01 AG026096 and P30 AG10133.

References

Anstey K, Christensen H. Education, Activity, Health, Blood Pressure and Apolipoprotein E as Predictors of Cognitive Change in Old Age: A Review. Gerontology. 2000; 46:163–177. [PubMed: 10754375]

Bravo G, Hebert R. Age- and education-specific reference valuates for the Mini-Mental and modified Mini-Mental state examinations derived from a non- Demented elderly population. International Journal of Geriatric Psychiatry. 1997; 12:1008–1018. [PubMed: 9395933]

Busch RM, Chapin JS. Review of normative data for common screening measures used to evaluate cognitive functioning in elderly individuals. The clinical Neuropsychologist. 2008; 22:620–650. [PubMed: 17853152]

Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. Journal of the American Medical Association. 1993; 269:2386–2391. [PubMed: 8479064]

Emsley CL, Gao S, Li Y, Liang C, Ji R, Hall KS, et al. Trace element levels in drinking water and cognitive function among elderly Chinese. American Journal of Epidemiology. 2000; 151(9):913–920. [PubMed: 10791564]

Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Global prevalence of dementia: a Delphi consensus study. Lancet. 2005; 366(9503):2112–2117. [PubMed: 16360788]

Ganguli M, Ratcliff G, Huff FJ, Belle S, Kancel MJ, Fischer L, et al. Effects of age, gender, and education on cognitive tests in a rural elderly community sample: Norms from the Monongahela Valley independent elders survey. Neuroepidemiology. 1991; 10:42–52. [PubMed: 2062416]

Ganguli M, Belle S, Ratcliff G, Seaberg E, Huff FJ, von der Porten K, Kuller LH. Sensitivity and specificity for dementia of population-based criteria for cognitive impairment: the MoVIES project. Journal of Gerontology. 1993; 48(4):152–161.

Gao S, Jin Y, Hall KS, Liang C, Unverzagt FW, Ji R, et al. Selenium level and cognitive function in rural elderly Chinese. Am J Epidemiol. 2007; 165:955–965. [PubMed: 17272290]

Hall KS, Hendrie HC, Brittain HM, et al. The development of a dementia screening interview in two distinct languages. International Journal of Methods in Psychiatric Research. 1993; 3:1–28.

Hall KS, Ogunniyi AO, Hendrie HC, et al. Community screening for dementia in Indianapolis and Ibadan, Nigeria. Neurobiology of Aging. 1994; 15:42.

Hall KS, Ogunniyi AO, Hendrie HC, Osuntokun BO, Hui Ô, SL, Musick MC, et al. A cross-cultural community based study of dementias: methods and performance of the survey instrument: Indianapolis, U.S.A. and Ibadan, Nigeria. International Journal of Methods in Psychiatric Research. 1996; 6:129–142.

Hall KS, Gao S, Emsley CL, Ogunniyi AO, Morgan O, Hendrie H. Community screening interview for dementia (CSI 'D'); performance in five disparate study sites. International Journal of Geriatric Psychiatry. 2000; 15(6):521–531. [PubMed: 10861918]

Hertzog, C. Research design in studies of aging and cognition. In: Birren, JE.; Schaie, KW., editors. Handbook of the Psychology of Aging. 4. San Diego: Academic Press; 1996. p. 24-35.

Hsieh SL, Tori CD. Normative data on cross-cultural neuropsychological tests obtained from Mandarin-speaking adults across the life span. Archives of Clinical Neuropsychology. 2007; 22(3):283–296. [PubMed: 17293080]

Isaacs B, Akhtar AJ. The set test: a rapid test of mental function in old people. Age Ageing. 1972; 1(4):222–226. [PubMed: 4669878]

Ishizaki J, Meguro K, Ambo H, Shimada M, Yamaguchi S, Hayasaka C, et al. A normative community-based study of Mini-Mental State in elderly adults: The effect ofage and educational level. Journal of Gerontology. 1998; 53B:359–363.

Kasai M, Ishizaki J, Ishii H, Yamaguchi S, Yamadori A, Meguro K. Normative data on Benton Visual Form Discrimination Test for older adults and impaired scores in Clinical Dementia Rating 0.5 participants: Community-based study. The Osaki–Tajiri Project. Psychiatry and Clinical Neurosciences. 2009; 63:9–16. [PubMed: 19154208]

Langa KM, Plassman BL, Wallace RB, Herzog AR, Heeringa SG, Ofstedal MB, et al. The Aging, Demographics, and Memory Study: study design and methods. Neuroepidemiology. 2005; 25:181–191. [PubMed: 16103729]

Lee DY, Lee KU, Lee JH, Kim KW, Jhoo JH, Kim SY, et al. A normative study of the CERAD neuropsychological assessment battery in the Korean elderly. Journal of the International Neuropsychological Society. 2004; 10:72–81. [PubMed: 14751009]

Lim ML, Collinson SL, Feng L, Ng TP. Cross-Cultural Application of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Performances of Elderly Chinese Singaporeans. Clinical Neuropsychologist. 2010; 24(5):811–826. [PubMed: 20560092]

Lin KN, Wang PN, Liu CY, Chen WT, Lee YC, Liu HC. Cutoff scores of the Cognitive Abilities Screening Instrument, Chinese version in screening of dementia. Dementia and Geriatric Cognitive Disorders. 2002; 14(4):176–182. [PubMed: 12411759]

Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology. 1989; 39(9):1159–1165. [PubMed: 2771064]

Østbye T, Krause KM, Norton MC, Tschanz J, Sanders L, Hayden K, et al. Ten dimensions of health and their relationships with overall self-reported health and survival in a predominately religiously active elderly population: The Cache County Memory Study. Journal of the American Geriatrics Society. 2006; 54:199–206. [PubMed: 16460369]

Parmenter BA, Testa SM, Schretlen DJ, Weinstock-Guttman B, Benedict RH. The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). J Int Neuropsychol Soc. 2010; 16(1):6–16. [PubMed: 19796441]

Schmidt R, Freidl W, Fazekas F, Reinhart B, Grieshofer P, Koch M, et al. The Mattis Dementia Rating Scale: Normative data from 1,001 healthy volunteers. Neurology. 1994; 44:964–966. [PubMed: 8190307]

Smerbeck AM, Parrish J, Yeh EA, Hoogs M, Krupp LB, Weinstock-Guttman B, Benedict RH. Regression-based pediatric norms for the brief visuospatial memory test: revised and the symbol digit modalities test. Clinical Neuropsychol. 2011; 25(3):402–12.

Snitz BE, Unverzagt FW, Chang C-C, Vander Bilt J, Gao S, Saxton J, et al. Effects of age, gender, education and race on two tests of language ability in community- based older adults. International Psychogeriatrics. 2009; 21(6):1051–1062. [PubMed: 19586563]

Sosa AL, Albanese E, Prince M, Acosta D, Ferri CP, Guerra M, et al. Population normative data for the 10/66 Dementia Research Group cognitive test battery from Latin America, India and China: a cross-sectional survey. Bmc Neurology. 2009; 9:48. [PubMed: 19709405]

Yamamoto K, Evans JD, Johnson KE, Unverzagt FW. Clinical utility of IU Token Test in the diagnosis of dementia. Journal of the International Neuropsychological Society. 2003; 9:316.

	A	B
1	**Subject Demographic**	
2	Age(years)	71
3	Gender(Male=0, Female=1)	1
4	Education(years)	6
5	Gender×Education	=B3*B4
6	**Observed CSID score**	26
7	**Multiple regression equation for CSID**	
8	Intercept	38.5114
9	Age(years)	-0.1738
10	Gender	-1.9156
11	Education(years)	0.1811
12	Gender×Education	0.3203
13	SD of residuals	3.1198
14	**CSID predicted score**	=B8+B2*B9+B3*B10+B4*B11+B5*B12
15	**CSID z-score**	=(B6-B14)/B13
16	**Percentile of CSID z-score**	=NORMSDIST(B15)

Figure.
Layout and formulas for calculating demographically-adjusted z-score and percentile equivalents.
Notes: CSID: the Community Screening Instrument for Dementia.
SD: standardized deviation.
NORMSDIST: statistical function in Excel used to calculate percentile for standardized normal distribution.

Table 1

Demographic and health history characteristics for the whole sample (n=1826) and by gender separately

	All (n=1826)	Male (n=837)	Female (n=989)	p-value
Age, M (SD)	71.9 (5.6)	71.5 (5.5)	72.2 (5.6)	0.0068
Years of schooling, M(SD)	1.6 (2.6)	2.8 (2.9)	0.6 (1.6)	<0.0001
0 years, n (%)	1226 (61.7)	290 (34.6)	836 (84.5)	
1–5 years, n (%)	498 (27.3)	377 (45.0)	121 (12.2)	
5+ years, n (%)	202 (11.1)	170 (20.3)	32 (3.2)	
Marital Status				<0.0001
Married, n (%)	1125 (61.6)	596 (71.2)	529 (53.5)	
Widowed, n (%)	679 (37.2)	224 (26.8)	455 (46.0)	
Other, n (%)	22 (1.2)	17 (2.0)	5 (0.5)	
Health History, n (%)				
Cancer	5 (0.3)	3 (0.4)	2 (0.2)	0.5245
Diabetes	41 (2.3)	10 (1.2)	31 (3.1)	0.0053
Hypertension	286 (15.7)	117 (14.0)	169 (17.1)	0.0702
Heart Attack	52 (2.8)	19 (2.3)	33 (3.3)	0.1721
Fracture	35 (1.9)	14 (1.7)	21 (2.1)	0.4847

Notes: p-value represents the difference between male and female for each demographic variable.

Table 2

Cumulative frequency distribution (and percentages) for test scores for the entire sample (n=1826).

	Global CSID	Memory			Language Animal Fluency	Executive IU Token
Raw Score		Word List Learning	Word List Recall	IU Story		
≥31	.	.	.	.	5 (100)	.
30	179 (100)	1 (100)	.	.	1 (99.7)	.
29	199 (90.2)		.	.	.	.
28	213 (79.3)		.	.	4 (99.7)	.
27	207 (67.6)		.	.	1 (99.5)	.
26	203 (56.3)	2 (99.9)	.	.	5 (99.4)	.
25	172 (45.2)	5 (99.8)	.	.	10 (99.1)	.
24	166 (35.8)	6 (99.6)	.	.	20 (98.6)	84 (100)
23	132 (26.7)	9 (99.2)	.	.	12 (97.5)	84 (95.4)
22	105 (19.4)	18 (98.7)	.	.	20 (96.8)	116 (90.8)
21	79 (13.7)	29 (97.8)	.	.	40 (95.7)	124 (84.4)
20	50 (9.4)	42 (96.2)	.	.	49 (93.5)	137 (77.6)
19	33 (6.6)	51 (93.9)	.	.	54 (90.9)	153 (70.1)
18	25 (4.8)	73 (91.1)	.	.	79 (87.9)	152 (61.8)
17	26 (3.5)	96 (87.1)	.	.	85 (83.6)	128 (53.4)
16	12 (2)	129 (81.8)	.	.	113 (78.9)	109 (46.4)
15	8 (1.4)	145 (74.8)	.	.	127 (72.7)	98 (40.4)
14	8 (0.9)	169 (66.8)	.	14 (100)	116 (65.8)	109 (35.1)
13	3 (0.5)	199 (57.6)	.	21 (99.2)	138 (59.4)	62 (29.1)
12	2 (0.3)	195 (46.7)	.	25 (98.1)	162 (51.9)	79 (25.7)
11	1 (0.2)	181 (36.0)	.	58 (96.7)	141 (43.0)	70 (21.4)
10	2 (0.2)	166 (26.1)	8 (100)	60 (93.5)	157 (35.3)	70 (17.5)
9	.	108 (17.0)	44 (99.6)	105 (90.2)	113 (26.7)	48 (13.7)
8	.	76 (11.1)	97 (97.2)	115 (84.3)	134 (20.5)	50 (11.1)
7	1 (0.1)	60 (6.9)	155 (91.8)	144 (78.0)	103 (13.1)	42 (8.3)
6	.	32 (3.6)	274 (83.4)	171 (70.0)	58 (7.5)	36 (6.0)
5	.	17 (1.9)	357 (68.3)	230 (60.6)	38 (4.3)	18 (4.1)
4	.	7 (0.9)	362 (48.8)	299 (47.8)	26 (2.2)	25 (3.1)

	Global CSID	Memory			Language Animal Fluency	Executive IU Token
Raw Score		Word List Learning	Word List Recall	IU Story		
3	.	4 (0.5)	313 (29.0)	283 (31.3)	8 (0.8)	15 (1.7)
2	.	2 (0.3)	141 (11.8)	222 (15.7)	7 (0.4)	11 (0.9)
1	.	1 (0.2)	40 (4.1)	61 (3.4)	.	2 (0.3)

Notes: CSID represents the Community Screening Instrument for Dementia.

Table 3

Parameter estimates and standard error estimates from multiple regression models regressing each cognitive score on age, gender, education and potential interactions.

Tests	Predictor	Estimate	Standard Error	P-value	Standardized Estimate	R^2	SD of residuals[#]
Global							
CSID	Intercept	38.5114	0.9826	<.0001			
	Age(years)	−0.1738	0.0134	<.0001	−0.2729		
	Gender[*]	−1.9156	0.1834	<.0001	−0.1971		
	Education(years)	0.1811	0.0375	<.0001	0.2540		
	Gender Education	0.3203	0.0733	<.0001	0.1143	0.2321	3.1198
Memory							
Word List Learning	Intercept	25.2878	1.1526	<.0001			
	Age(years)	−0.1711	0.0157	<.0001	−0.2432		
	Gender[*]	−0.6967	0.1920	0.0003	−0.0884		
	Education(years)	0.2775	0.0383	<.0001	0.1800	0.1336	3.6595
Word List Recall	Intercept	10.6978	0.5729	<.0001			
	Age(years)	−0.0843	0.0078	<.0001	−0.2434		
	Gender[*]	−0.3220	0.0954	0.0008	−0.0829		
	Education(years)	0.1155	0.0190	<.0001	0.1521	0.1178	1.8191
IU Story	Intercept	10.5936	0.8660	<.0001			
	Age(years)	−0.0792	0.0119	<.0001	−0.1536		
	Education	0.2574	0.0258	<.0001	0.2278	0.0894	2.7491
Language							
Animal Fluency Test	Intercept	27.3504	1.3889	<.0001			
	Age(years)	−0.1988	0.0189	<.0001	−0.2284		
	Gender[*]	−1.6717	0.2314	<.0001	−0.1714		
	Education(years)	0.3915	0.0461	<.0001	0.2052	0.1785	4.4099
Executive							
IU Token Test	Intercept	35.0553	1.4937	<.0001			
	Age(years)	−0.2639	0.0203	<.0001	−0.2762		
	Gender[*]	−1.5906	0.2784	<.0001	−0.0917		

Tests	Predictor	Estimate	Standard Error	P-value	Standardized Estimate	R^2	SD of residuals[#]
	Education(years)	0.4320	0.0570	<.0001	0.3053		
	Gender Education	0.3832	0.1113	0.0006	0.0912	0.2144	4.7348

Notes:

[*] Gender is coded as female=1 versus male=0;

SD of residuals[#]: the root of mean square of errors (RMSE).